ENTRE DE ROCQUENCOURT

Institut National
de Recherche
en Informatique
et en Automatique

Domaine de Voluceau
Rocquencourt
B.P.105
78153 Le Chesnay Cedex
France
Tél. : (1) 39 63 55 11

Rapports de Recherche

N° 482

UNE MÉTHODE A MÉTRIQUE VARIABLE RÉDUITE EN OPTIMISATION AVEC CONTRAINTES D'ÉGALITÉ NON LINÉAIRES

Jean Charles GILBERT

Janvier 1986

Octobre 1985

UNE METHODE A METRIQUE VARIABLE REDUITE EN OPTIMISATION AVEC CONTRAINTES D'EGALITE NON LINEAIRES

Jean Charles GILBERT

boursier scientifique et technique
de la Commission des Communautés Européennes

(adresse actuelle)
25, rue Albert
F - 75013 Paris

Soumis à la revue
MODELISATION MATHEMATIQUE ET ANALYSE NUMERIQUE

UNE METHODE A METRIQUE VARIABLE REDUITE EN

OPTIMISATION AVEC CONTRAINTES D'EGALITE NON LINEAIRES

Jean Charles GILBERT

Résumé: En optimisation dans $\mathbb{R}^n$ avec m contraintes d'égalité non linéaires, on propose une modification de la méthode séquentielle quadratique de Han où il ne faut mettre à jour qu'une matrice d'ordre n-m, plutôt que d'ordre n. La convergence est Q-superlinéaire (en un pas). On introduit une méthode globalisant l'algorithme local consistant à déterminer un pas sur un arc de courbe de manière à faire décroître une fonctionnelle pénalisée exacte. Asymptotiquement le pas vaut un. On traite également de la mise à jour de la matrice réduite en considérant le schéma BFGS dans le cadre de la méthode globale.

Abstract: In optimization in $\mathbb{R}^n$ with m nonlinear equality constraints, a modification of the sequential quadratic method of Han is proposed in which a matrix of order n-m instead of order n is updated. The convergence is Q-superlinear (in one step). We introduce a method globalizing the local algorithm which consists in determining a step on an arc in order to decrease an exact penalty function. Asymptoticaly, the step is equal to one. We also deal with the matrix update by considering the BFGS formula within the framework of the global method.

J'aimerais remercier J.P. Puel et J.F. Bonnans pour leurs encouragements, leurs remarques et suggestions ainsi que pour avoir lu avec attention le manuscrit.

TABLE DES MATIERES

I - INTRODUCTION

Soient Ω un ouvert de $\mathbb{R}^n$, f et c deux applications définies sur Ω à valeurs dans $\mathbb{R}$ et $\mathbb{R}^m$ $(m < n)$ respectivement. On considère le problème d'optimisation avec contraintes d'égalité suivant:

$$\min \{ f(x) : x \in \Omega , c(x) = 0 \} \tag{1.1}$$

Soit x^* un point satisfaisant aux conditions nécessaires d'optimalité du premier ordre de (1.1). On s'intéresse à un algorithme à métrique variable pour ce problème. C'est-à-dire que l'on cherche à construire une suite de points x_k convergeant vers x^* et une suite de matrices définissant des métriques variables qui sont sensées converger, dans un certain sens, vers une matrice dont le calcul ferait intervenir les dérivées secondes de f et c au point optimal x^*, information supposée non accessible. Ces méthodes sont particulièrement intéressantes du fait qu'en général leur ordre de convergence est Q-superlinéaire.

Un algorithme de ce type consisterait à utiliser une méthode de quasi-Newton pour résoudre le système d'optimalité du premier ordre de (1.1):

$$\left| \begin{array}{l} f(x^*) + c'(x^*)^T \lambda^* = 0 \\ c(x^*) = 0 \end{array} \right. \tag{1.2}$$

dans lequel $\lambda^* \in \mathbb{R}^m$ est le multiplicateur de Lagrange associé aux contraintes en x^*. Cela conduirait à la mise à jour d'une matrice d'<u>ordre (n+m)</u>, approximation de la jacobienne du système (1.2) au point optimal (x^*,λ^*). Par exemple, l'algorithme de Garcia, Mangasarian (1976) met à jour une matrice ayant cette propriété.

La méthode séquentielle quadratique (MSQ) proposée par Han (1976) améliore la méthode précédente en ce qui concerne la taille de la matrice à mettre à jour. Elle consiste en la résolution d'une suite de problèmes quadratiques avec contraintes linéaires. Etant donné x_k, on calcule x_{k+1} comme solution du problème:

$$\left|\begin{array}{l} \min\limits_{x} \{ f'(x_k).(x-x_k) + \frac{1}{2} M_k.(x-x_k)^2 \} \\ c(x_k) + c'(x_k).(x-x_k) = 0 \end{array}\right. \qquad (1.3)$$

A chaque itération, M_k, qui est d'ordre n, est mise à jour afin d'approcher L^*, le hessien du lagrangien l au point optimal.

$$l(x,\lambda) := f(x) + c(x)^T \lambda \qquad (1.4)$$

$$L(x,\lambda) := \partial_{xx}\, l(x,\lambda) \qquad (1.5)$$

$$L^* := L(x^*,\lambda^*) \qquad (1.6)$$

Dans ce travail, on propose une méthode Q-superlinéairement convergeante en un pas (c'est-à-dire $||x_{k+1}-x^*||/||x_k-x^*||$ converge vers 0) dans laquelle on ne met à jour qu'une matrice d'ordre (n-m). Cette méthode est alors bien adaptée aux problèmes où n est grand et (n-m) petit. Dans cette classe se situent par exemple les problèmes d'identification paramétrique dans les équations aux dérivées partielles. Si on adopte une discrétisation des équations d'état (les contraintes) par une méthode d'éléments finis, m est grand, disons de l'ordre de 1000. Tandis que le nombre (n-m) de paramètres identifiables est faible: 2 ou 3 dans l'exemple de Blum, Gilbert, Thooris (1985). Dans cet exemple, la méthode présentée est utilisable alors que la MSQ ne l'était pas.

Plusieurs travaux dans cette direction ont déjà été réalisés. Powell (1978) a montré que la suite (x_k) générée par la MSQ était indépendante d'une sous-matrice d'ordre m de M_k. Il obtenait également une condition suffisante de convergence Q-superlinéaire en deux pas pour la MSQ (c'est-à-dire $||x_{k+1}-x^*||/||x_{k-1}-x^*||$ converge vers 0) portant sur une sous-matrice d'ordre (n-m) de M_k. Toutefois, il ne donne pas de méthode pour mettre à jour cette "petite" sous-matrice. Dans la section III, nous donnons une condition nécessaire et suffisante de convergence Q-superlinéaire pour la MSQ qui s'apparente à la condition de Dennis, Moré (1974) pour les méthodes à métriques variables en optimisation sans contrainte. Cette condition et le contre-exemple qui suit montrent qu'en général il n'est pas possible d'obtenir la convergence Q-superlinéaire (en un pas) de la MSQ en ne mettant à jour qu'une matrice "réduite" (d'ordre n-m plutôt que d'ordre n).

Cette remarque nous a conduit à modifier la MSQ afin qu'avec la mise à jour d'une matrice réduite, notée G_k, on garde la convergence Q-superlinéaire (section IV). Une itération de la MSQ peut se décomposer en deux phases: une phase de restauration des contraintes et une phase de minimisation du critère f. Dans l'algorithme proposé, la phase de minimisation ne demande que le calcul d'un gradient réduit. Ce calcul se fait en linéarisant les contraintes <u>après</u> la phase de restauration. Ce point est essentiel. En effet, deux autres algorithmes ne mettant à jour que des matrices réduites on été proposés. Dans la méthode de Gabay (1982,b), la linéarisation des contraintes et le calcul du gradient réduit se font <u>avant</u> la restauration et la convergence du procédé est Q-superlinéaire mais en deux pas. Dans la méthode proposée par Coleman, Conn (1982,a), la restauration suit la phase de minimisation: la convergence est Q-superlinéaire en deux pas.

La méthode présentée sépare clairement les deux problèmes posés par (1.1). D'abord, en un point x, on cherche à satisfaire aux contraintes , $c(x^*) = 0$, par une phase de restauration qui s'apparente à une itération de Newton: la linéarisation des contraintes se faisant après la phase de restauration, on ne dispose pas de la jacobienne en x mais d'un opérateur proche de celle-ci, par exemple, la jacobienne calculée à l'itération précédente. Ensuite, étant passé de x à y par la phase de restauration des contraintes, on cherche à minimiser f par une méthode de quasi-Newton réduite par le plan tangent à la variété $c^{-1}(c(y))$. C'est cette réduction qui permet de ne mettre à jour qu'une matrice d'ordre n-m, exactement comme si on cherchait à minimiser f en présence de m contraintes d'égalité linéaires.

On peut aussi voir la méthode présentée comme faisant le lien entre la MSQ et la méthode de Gradient Réduit Généralisée (GRG). En effet, elle est issue de la MSQ par modification de son déplacement. Elle peut également être vue comme une méthode de GRG où la phase de restauration ne comprend qu'un seul pas et la phase de minimisation se fait dans le plan tangent à une variété "parallèle" à $c^{-1}(0)$.

En ce qui concerne la globalisation de la méthode (section V), on s'y prend de manière classique, par pénalisation exacte du critère f:

$$\Theta_p(x) = f(x) + p\ ||c(x)||_1 \tag{1.7}$$

où p est un réel positif. Si on choisit p assez grand, les points stationnaires de (1.1) sont des points stationnaires de Θ_p et peuvent alors être recherchés par minimisation de Θ_p. Comme dans la MSQ, on cherche les directions de descente de Θ_p en un point par les déplacements fournis par l'algorithme décrit. Soit y_k, le point obtenu à partir de x_k après la phase de restauration des contraintes. Du fait que la restauration se fait en x_k et que la linéarisation des contraintes se fait en y_k, on ne peut en général pas dire si $(x_{k+1}-x_k)$ ou $(y_{k+1}-y_k)$ sont des directions de descente de Θ_p en x_k ou y_k respectivement. Toutefois, $q_k=x_{k+1}-y_k$ est une direction de descente en y_k si et seulement si q_k est non nul. Dès lors, on définit en y_k un arc parabolique de descente de Θ_p:

$$y_k(\rho) = y_k + \rho\ (x_{k+1}-y_k) + \rho^2\ (y_{k+1}-x_{k+1}) \qquad (1.8)$$

On montre qu'on fait décroître Θ_p si on prend ρ assez petit. La recherche parabolique se fera donc à partir de y_k et non de x_k. Un critère (C) joint à une règle de type Armijo permet de déterminer ρ_k à chaque itération. Ceci conduit à un théorème de convergence globale. De plus, asymptotiquement, le pas unité ($\rho_k=1$) est admissible pour le critère (C). Dès lors, la recherche parabolique (1.8) n'empêche pas la convergence Q-superlinéaire locale: on a un bon recollement des méthodes utilisées loin du point optimal et proche de celui-ci.

Il reste à traiter le problème de la mise à jour de la matrice réduite G_k (section VI). L'étude se fera dans le cadre de la méthode globale introduite à la section V, c'est-à-dire que l'on ne supposera pas ρ_k égal à 1. Localement, la matrice G_k apparaît comme une approximation de G^*, réduction au plan tangent aux contraintes du hessien du lagrangien pris au point optimal (1.6). Pour l'algorithme qu'il étudie, Gabay (1982,b) propose une équation de quasi-Newton par analogie avec celle obtenue dans les méthodes de type gradient-restauration (Gabay (1982,a)), mais n'en montre pas la validité dans son nouveau cadre. Nocedal, Overton (1985) étudient la mise à jour de la matrice réduite de l'algorithme de Gabay (1982,b) dans un cadre local. Coleman, Conn (1984) proposent pour leur algorithme une équation de quasi-Newton nécessitant deux linéarisations des contraintes par itération. D'une part, nous reprenons la méthode de Coleman-Conn en simplifiant leur équation de quasi-Newton (algorithme ALG1), et d'autre part, nous proposons une méthode ne nécessitant qu'une seule linéarisation des contraintes par itération (algorithme ALG2). Dans cette dernière méthode, un critère décide à chaque itération de l'opportunité d'une mise à jour de G_k. Pour ces deux algorithmes, on obtient la convergence Q-superlinéaire en un pas.

II - NOTATIONS ET HYPOTHESES

II.1 - Définitions et notations.

Soient Ω un ouvert de $\mathbb{R}^n$ et

$f: \Omega \to \mathbb{R}$

$c: \Omega \to \mathbb{R}^m$, $m < n$.

On considère le problème d'optimisation

$$\min \{ f(x) : x \in \Omega,\ c(x) = 0 \} \tag{2.1}$$

On définit le lagrangien l sur $\Omega \times \mathbb{R}^m$ par

$$l(x,\lambda) = f(x) + c(x)^T \lambda \tag{2.2}$$

et on désigne par $L(x,\lambda)$ son hessien:

$$L(x,\lambda) = \partial_{xx} l(x,\lambda) = f''(x) + \lambda \circ c''(x) \tag{2.3}$$

$$\lambda \circ c''(x) := \sum_{i=1}^{m} \lambda_i\, c_i''(x), \text{ une matrice d'ordre } n. \tag{2.4}$$

On notera indifféremment par $||.||$ les normes euclidiennes de $\mathbb{R}$, $\mathbb{R}^{n-m}$, $\mathbb{R}^m$ et $\mathbb{R}^n$ ainsi que les normes matricielles subordonnées à ces normes vectorielles.

Soient (x_k) une suite de points de Ω, φ et ψ deux applications définies sur Ω à valeurs dans un espace normé (par $||.||$) et $\mathbb{R}^+$ respectivement. On notera $\varphi(x_k) = O(\psi(x_k))$ (grand O) s'il existe une constante positive C telle que pour tout indice k on ait:

$$||\varphi(x_k)|| \leq C\ \psi(x_k)$$

On notera $\varphi(x_k) = o(\psi(x_k))$ (petit o) si quelque soit $\varepsilon > 0$, il existe un indice k_ε tel que pour tout indice k plus grand que k_ε on ait:

$$||\varphi(x_k)|| \leq \varepsilon\ \psi(x_k)$$

On dira qu'une suite de matrices carrées (M_k) est uniformément non singulière s'il existe une constante positive C telle que l'on ait:

$$||\ M_k\ v\ || \geq C\ ||\ v\ ||$$

pour tout indice k et tout vecteur v. On dira que la suite (M_k) est uniformément définie positive s'il existe une constante positive C telle que l'on ait:

$$(M_k\, v\, ,\, v) \geq C\, ||v||^2$$

pour tout indice k et tout vecteur v.

II.2 - Changement de coordonnées.

On se place dans le cadre défini par Gabay (1982,a). Soient $x \in \Omega$ et $A_x := c'(x)$ qui est une matrice m x n. Supposons que A_x soit surjective. Alors, on peut choisir, de façon non unique, un inverse à droite A_x^- de A_x:

$$A_x\, A_x^- = I_m \tag{2.5}$$

$$N(A_x) \wedge R(A_x^-) = \{0\} \tag{2.6}$$

La relation (2.6) se déduit de (2.5). D'après (2.6), le noyau de A_x, $N(A_x)$, et l'image de A_x^-, $R(A_x^-)$, sont des sous-espaces complémentaires non nécessairement orthogonaux, de dimension (n-m) et m respectivement. A_x^- est injective. On peut également choisir, de façon non unique, une matrice Z_x de dimension (n-m) x n, telle que:

$$S_x := \begin{bmatrix} A_x \\ Z_x \end{bmatrix} \text{ non singulière } \iff N(A_x) \wedge N(Z_x) = \{0\} \tag{2.7}$$

$$N(Z_x) = R(A_x^-) \tag{2.8}$$

En effet, il suffit que les (n-m) lignes de Z_x forment une base du complémentaire orthogonal de $R(A_x^-)$. Alors on a (2.8), puis (2.7) grâce à (2.6). D'après (2.7), Z_x est surjective et la relation (2.8) montre que l'on a:

$$Z_x\, A_x^- = 0 \tag{2.9}$$

Pour ces choix de A_x^- et Z_x, il existe une unique matrice Z_x^-, de dimension n x (n-m), vérifiant

$$Z_x\, Z_x^- = I_{n-m} \tag{2.10}$$

$$N(A_x) = R(Z_x^-) \tag{2.11}$$

En effet, soit M_x une matrice de dimension n x (n-m) dont les colonnes forment une base de $N(A_x)$. Alors, on a $R(M_x) = N(A_x)$ et d'après (2.7), $N(Z_x) \wedge R(M_x) = \{0\}$. Ceci montre que $Z_x M_x$ est non singulière. Alors, $Z_x^- := M_x(Z_x M_x)^{-1}$ vérifie (2.10) et (2.11). L'unicité de Z_x^- vient de ce que si Z_1^- et Z_2^- sont deux matrices jouant le rôle de Z_x^- dans (2.10) et (2.11), on a $\mathrm{col}(Z_1^- - Z_2^-) \in N(Z_x) \wedge N(A_x)$. D'après (2.7), cela montre que $Z_1^- = Z_2^-$. D'après (2.10), Z_x^- est un inverse à droite de Z_x et d'après (2.11) on a:

$$A_x\, Z_x^- = 0 \tag{2.12}$$

Grâce à (2.5), (2.9), (2.10) et (2.12), on vérifie que l'inverse de S_x est donné par

$$S_x^{-1} = [\, A_x^- \quad Z_x^- \,] \tag{2.13}$$

L'identité $I_n = S_x^{-1}\, S_x$ s'écrit:

$$I_n = A_x^- A_x + Z_x^- Z_x \tag{2.14}$$

D'après (2.13) et la régularité de S_x^{-1}, on voit que $\mathbb{R}^n$ peut se décomposer en deux sous-espaces complémentaires mais non nécessairement orthogonaux: $R(A_x^-)$ et $R(Z_x^-)$. La relation (2.14) permet de décomposer un vecteur ξ de $\mathbb{R}^n$ en ses composantes dans $N(A_x) = R(Z_x^-)$ et $N(Z_x) = R(A_x^-)$:

$$\xi = A_x^- A_x\, \xi + Z_x^- Z_x\, \xi$$

$A_x\xi$ sont les coordonnées de ξ dans $N(Z_x)$ et $Z_x\xi$ les coordonnées de ξ dans $N(A_x)$.

Pour fixer les idées, considérons deux exmples. Dans le cadre des problèmes de contrôle, on partitionne A_x en:

$$A_x = \begin{bmatrix} B_x & D_x \end{bmatrix} \tag{2.15}$$

avec B_x d'ordre m, non singulière et D_x de dimension m x (n-m). On prend pour A_x^-, l'<u>inverse à droite partitionné</u> de A_x:

$$A_x^- = \begin{bmatrix} B_x^{-1} \\ 0 \end{bmatrix} \tag{2.16}$$

Et pour Z_x, on choisit la matrice (n-m) x n:

$$Z_x = \begin{bmatrix} 0 & I_{n-m} \end{bmatrix} \tag{2.17}$$

Alors, Z_x^- est déterminé et s'écrit:

$$Z_x^- = \begin{bmatrix} -\,B_x^{-1}\,D_x \\ I_{n-m} \end{bmatrix} \tag{2.18}$$

On peut également prendre pour A_x^-, le <u>pseúdo-inverse de Penrose</u>:

$$A_x^- := A_x^T\,(A_x\,A_x^T)^{-1} \tag{2.19}$$

Alors, on a les relations

$$N(A_x)^{\perp} = R(A_x^T) = R(A_x^-) \tag{2.20}$$

On prend pour lignes de Z_x une base de $R(A_x^-)^{\perp} = N(A_x)$. Alors Z_x^- est déterminé et s'écrit (on prend $M_x = Z_x^T$ dans le raisonnement qui suit (2.11)):

$$Z_x^- = Z_x^T\,(Z_x\,Z_x^T)^{-1} \tag{2.21}$$

C'est donc le pseudo-inverse de Penrose de Z_x. On a les relations:

$$N(A_x) = R(Z_x^T) = R(Z_x^-) \tag{2.22}$$

II.3 - Hypothèses générales.

II.3.1 - Hypothèse H_0.

On suppose que f et c sont deux fois continûment différentiables (C^2) sur Ω et que leurs dérivées secondes f'' et c'' sont bornées sur Ω. On suppose également que c est une <u>submersion</u> sur Ω. C'est-à-dire que pour tout x dans Ω, la matrice m x n

$$A(x) := c'(x)$$

est surjective. La surjectivité de c' en x^* est une hypothèse de qualification des contraintes (Ciarlet (1982)). Par continuité, elle implique la surjectivité de c' dans un voisinage de x^*. Dans le cadre d'une étude locale, il suffit donc de supposer $A(x^*)$ surjective et Ω assez petit.

On notera A, A^-, Z, Z^- les opérateurs $A(x^*)$, $A(x^*)^-$, $Z(x^*)$, $Z(x^*)^-$.

II.3.2 - Hypothèse H_1: conditions d'optimalité.

On suppose qu'il existe un couple (x^*,λ^*) dans $\Omega \times \mathbb{R}^m$ vérifiant les conditions nécessaires d'optimalité du premier ordre de (2.1):

$$\left| \begin{array}{l} f'(x^*) + A^T \lambda^* = 0 \\ c(x^*) = 0 \end{array} \right. \tag{2.23}$$

x^* est aussi appelé point de Kuhn-Tucker.

En multipliant la première équation de (2.23) par Z^{-T} et par A^{-T} et en tenant compte de (2.12) et de (2.5), on obtient les relations:

$$Z^{-T} f'(x^*) = 0 \tag{2.24}$$

$$\lambda^* = - A^{-T} f'(x^*) \tag{2.25}$$

L'équation (2.24) exprime que le gradient de f en x^* est orthogonal au plan tangent aux contraintes en x^*: $N(A) = R(Z^-)$.

II.3.3 - Hypothèse H_2: non singularité du hessien réduit.

On note $L^* := L(x^*,\lambda^*)$. On suppose que L^* restreinte au plan tangent aux contraintes en x^* est non singulière, c'est-à-dire que la matrice d'ordre $(n-m)$

$$G^* := Z^{-T} L^* Z^- \tag{2.26}$$

est non singulière. Remarquons qu'on ne demande pas la définie positivité de L^* dans le plan tangent, hypothèse souvent faite et qui fait partie des conditions suffisantes d'optimalité du second ordre. Supposer G^* seulement non singulière permet de prendre en compte les problèmes de recherche de points stationnaires du problème (1.1). On note l'inverse de G^* par:

$$H^* := (G^*)^{-1} \tag{2.27}$$

II.3.4 - Hypothèse H_3.

Si c est de classe C^σ ($\sigma \geq 2$) sur Ω, on suppose que l'on choisit l'application

$$x \longrightarrow (A_x, A_x^-, Z_x, Z_x^-)$$

de classe $C^{\sigma-1}$, à dérivées t-ième ($0 \leq t \leq \sigma-1$) bornées sur Ω (on restreint Ω si nécessaire). Les matrices A_x^T, A_x^-, Z_x^T, Z_x^- étant injectives et continues en x, on suppose qu'on a les propriétés d'uniforme injectivité suivantes. Il existe une constante positive C telle que pour tout x dans Ω, pour tout v dans $\mathbb{R}^m$ et pour tout w dans $\mathbb{R}^{n-m}$ on ait:

$$||A_x^T v|| \geq C\ ||v||, \qquad ||A_x^- v|| \geq C\ ||v||,$$

$$||Z_x^T w|| \geq C\ ||w||, \qquad ||Z_x^- w|| \geq C\ ||w||.$$

II.4 - **Le gradient réduit.**

Soient $x \in \Omega$ et $V_x := c^{-1}(c(x))$. Si c est une submersion de classe C^2, V_x est une variété différentiable de classe C^2 (Milnor (1965)). On munit V_x d'une structure riemannienne en définissant sur son plan tangent $T_x = R(Z_x^-)$ en x un produit scalaire:

$$\gamma_x\ (\ Z_x^- h\ ,\ Z_x^- g\ ) = (\ h\ ,\ g\ )_{\mathbb{R}^{n-m}} \tag{2.28}$$

où $Z_x^- h$ et $Z_x^- g$ sont dans T_x. On peut alors calculer le gradient de f sur V_x, noté $\nabla_x f$, et défini par:

$$\nabla_x f \ \in\ R(Z_x^-)$$

$$\gamma_x\ (\ \nabla_x f\ ,\ Z_x^- g\ ) = f'(x).(Z_x^- g) \quad \text{pour tout } g \in \mathbb{R}^{n-m}$$

où, f'(x) est la dérivée de f en x dans $\mathbb{R}^n$. Avec (2.28), on trouve:

$$\nabla_x f = Z_x^{-} Z_x^{-T} f'(x) \quad \varepsilon \quad \mathbb{R}^n$$

On dira que

$$g_x := g(x) := Z_x^{-T} f'(x) \quad \varepsilon \quad \mathbb{R}^{n-m} \qquad (2.29)$$

est le <u>gradient réduit</u> de f en x sur V_x. La condition d'optimalité (2.24) exprime que le gradient réduit est nul au point optimal x^*.

On dira que la matrice G^*, définie en (2.26), est le <u>hessien réduit du lagrangien</u> au point optimal x^*.

Le gradient réduit g(x), donné par (2.29), va jouer ici le rôle du gradient en optimisation sans contrainte. Sous les hypothèses (H_0) et (H_3), l'application

$$x \longrightarrow g(x) = Z_x^{-T} f'(x) \qquad (2.30)$$

est C^1 sur Ω. On calcule aisément la dérivée du gradient réduit au point optimal x^* comme suit (voir Nocedal, Overton (1985) qui donnent en référence des travaux de Goodman et de Stoer):

$$g'(x^*) = (Z_x^{-T} (f'(x) + A_x^T \lambda^*))' (x^*) = Z^{-T} L^* \qquad (2.31)$$

où on a tenu compte de (2.12) et de première équation de (2.23).

III - UNE CNS DE CONVERGENCE SUPERLINEAIRE POUR LA M.S.Q.

III.1 - La méthode séquentielle quadratique.

III.1.1 - Définition.

On utilise l'indice k pour les grandeurs évaluées en $x_k \in \Omega$:

$$A_k := A(x_k) = c'(x_k)$$

$$c_k := c(x_k), \text{ etc } \ldots$$

La méthode séquentielle quadratique (MSQ) consiste à générer la suite (x_k) à partir de $x_0 \in \Omega$ comme suit:

$$x_{k+1} = x_k + d_k \tag{3.1}$$

où d_k, le déplacement, est solution du problème quadratique:

$$\left| \begin{array}{l} \min\limits_{d} \{ f'(x_k).d \ + \frac{1}{2} M_k.d^2 \} \\ c_k + A_k \, d = 0 \end{array} \right. \tag{3.2}$$

La matrice M_k, d'ordre n, est mise à jour à chaque itération sur base d'une équation de quasi-Newton. Localement, on cherche à obtenir M_k aussi proche que possible du hessien L^* (voir Han (1976), Gabay (1982,b)).

III.1.2 - Hypothèses sur M_k (H_4).

On suppose:

$$(H_4) \left| \begin{array}{ll} . \ M_k \text{ symétrique} & (3.3) \\ . \ (M_k) \text{ bornée} & (3.4) \\ . \ M_k \text{ uniformément non singulière dans le plan } N(A_k) & (3.5) \end{array} \right.$$

On note G_k, la réduction de M_k au plan $N(A_k)$, tangent à $c^{-1}(c_k)$ en x_k:

$$G_k := Z_k^{-T} M_k Z_k^{-} \tag{3.6}$$

L'hypothèse (3.5) signifie que G_k est uniformément non singulière. Dans ce cas $H_k := G_k^{-1}$ est bien définie et la suite (H_k) est bornée. Remarquons que l'on ne demande pas que M_k soit non singulière comme dans Powell (1978) ou Boggs, Tolle, Wang (1982), ni même que G_k soit définie positive comme dans les articles précédents et dans Gabay (1982,b).

III.1.3 - Lemme.

> SI . c est différentiable en x
> . $A_x := c'(x)$ est surjective
> . M_x une matrice d'ordre n telle que sa réduction à $N(A_x)$, $Z_x^{-T} M_x Z_x^-$, soit non singulière
> (Z_x^- choisie comme en II.2)
>
> ALORS la matrice
> $$\begin{bmatrix} M_x & A_x^T \\ A_x & 0 \end{bmatrix}$$
> est non singulière.

Preuve.
Soit $(y,\lambda) \in \mathbb{R}^n \times \mathbb{R}^m$ tel que

$$M_x\, y + A_x^T \lambda = 0 \tag{3.7}$$

$$A_x\, y \qquad = 0 \tag{3.8}$$

Il suffit de montrer que (y,λ) est nul. D'après (3.8), $y \in N(A_x) = R(Z_x^-)$. Soit alors $h \in \mathbb{R}^{n-m}$ tel que

$$y = Z_x^- h$$

En multipliant (3.7) à gauche par Z_x^{-T} et en tenant compte de (2.12), on obtient

$$Z_x^{-T} M_x Z_x^- h = 0$$

La non singularité de $Z_x^{-T} M_x Z_x^-$ montre que h est nul; et donc aussi y. Alors (3.7) et l'injectivité de A_x^T montrent que λ est nul. ●●●

III.1.4 - Le déplacement.

Les équations d'optimalité du premier ordre de (3.2) s'écrivent

$$\left|\begin{array}{l} f'(x_k) + M_k\, d_k + A_k^T\, \lambda_{k+1} = 0 \\ c_k + A_k\, d_k = 0 \end{array}\right. \tag{3.9}$$

Les hypothèses (H_0) et (H_4) et le lemme III.1.3 montrent que l'on peut résoudre (3.9) en (d_k, λ_{k+1}). Comme dans Gabay (1982,b), on trouve:

$$d_k = - A_k^- c_k - Z_k^- H_k [\, g_k - Z_k^{-T} M_k A_k^- c_k \,] \tag{3.10}$$

Ce qui s'écrit également:

$$d_k = - Z_k^- H_k\, g_k - [\, I - Z_k^- H_k Z_k^{-T} M_k \,] A_k^- c_k \tag{3.11}$$

Dans (3.10), $(-A_k^- c_k) \in N(Z_k)$ peut être vu comme une phase de restauration des contraintes (une itération de Newton) et le second terme du second membre, qui est dans $N(A_k)$, comme une phase de minimisation du critère f. On y distingue le gradient réduit de f en x_k: $g_k = Z_k^{-T} f'(x_k)$. Mentionons que le déplacement de la méthode de Gabay (1982,b) se déduit de (3.10) en négligeant une partie du déplacement tangent:

$$d_k = - A_k^- c_k - Z_k^- H_k\, g_k \tag{3.12}$$

Si d_k est nul, (3.10) multipliée par A et par Z montre que $c_k=0$ et que $g_k=0$. C'est-à-dire que x_k vérifie les conditions d'optimalité du premier ordre (2.23). D'après l'hypothèse (H_2) et le lemme III.1.3, la jacobienne de (2.23) est non singulière. x^* est donc solution unique de (2.23) dans un voisinage de x^*. Dans le cadre d'une étude locale, on peut supposer que Ω est contenu dans ce voisinage. Dès lors, $d_k=0$ implique $x_k=x^*$. On supposera que cette situation ne se présente pas (arrêt de l'algorithme). C'est-à-dire que pour tout indice k on aura:

$$x_k \neq x^*; \quad x_{k+1} \neq x_k; \quad d_k \neq 0.$$

III.2 - <u>CNS de convergence Q-superlinéaire.</u>

III.2.1 - <u>Théorème.</u>

SI . (H_0), (H_1), (H_2) et (H_3)

. $(x_k) \subset \Omega$ générée par la MSQ avec (H_4)

. $x_k \underset{\infty}{\longrightarrow} x^*$

ALORS les clauses suivantes sont équivalentes

(i) la convergence de x_k vers x^* est Q-superlinéaire.

(ii) $Z^{-T} (M_k - L^*) (x_k - x^*) = o(||x_k - x^*||)$ (3.13)

(iii) $Z^{-T} (M_k - L^*) d_k = o(||x_k - x^*||)$ (3.14)

(iv) $g(x_{k+1}) = o(||x_k - x^*||)$ (3.15)

Ce théorème est à rapprocher de théorème 2.2 de Dennis, Moré (1974), établi pour les méthodes de quasi-Newton sans contrainte. Boggs, Tolle, Wang (1982) ont montré l'équivalence entre (i) et (iii) mais en supposant M_k non singulière et uniformément définie positive dans le plan $N(A_k)$, alors que nous ne supposons que la non singularité de M_k dans ce même plan (hypothèse H_4). De plus, ils supposaient la convergence linéaire de x_k vers x^*. Leur CNS de convergence s'écrit:

$$P^T (M_k - L^*) d_k = o(||x_k - x^*||) \tag{3.16}$$

où P est le projecteur sur N(A) parrallèlement à N(Z):

$$P = I - A^- A = Z^- Z$$

lorsqu'on choisit pour A^- le pseudo-inverse de Penrose (2.19). Du fait de l'injectivité de Z^-, (3.16) est équivalent à (3.14) pour ce choix particulier de A^-. Nocedal, Overton (1985) ont montré de façon élégante l'équivalence entre (i) et (ii).

Sous les hypothèses du théorème et notamment l'uniforme non singularité de H_k (obtenue grâce à (3.4) et (H_3)), on peut montrer que

$$||d_k|| \sim ||x_k - x^*|| \qquad (3.17)$$

où le signe $\sim$ signifie qu'il existe deux constantes positives c_1 et c_2 indépendantes de k telles que:

$$c_1 \; ||d_k|| \leq ||x_k - x^*|| \leq c_2 \; ||d_k||$$

Pour voir cela, on consultera Powell (1978) ou Gabay (1982,b). On pourra également utiliser l'argument présenté ultérieurement pour démontrer (3.17) dans le cadre de la nouvelle méthode. L'équivalence (3.17) est aussi une conséquence directe de la convergence Q-superlinéaire de (x_k). Dès lors, dans (3.13), (3.14) et (3.15), on pourra remplacer $o(||x_k-x^*||)$ par $o(||d_k||)$.

Le sujet ayant été abondamment traité, nous ne donnerons pas de preuve du théorème. Celle-ci peut se trouver dans Gilbert (?). On peut également s'inspirer de la preuve de la CNS de convergence Q-superlinéaire donnée pour la méthode présentée à la section IV.

III.3 - Un contre-exemple.

Les clauses (ii) et (iii) du théorème précédent indiquent que pour obtenir la convergence Q-superlinéaire, il faudra en général que la mise à jour de M_k soit telle que la partie $Z_k^{-T}M_k$ de M_k approche bien $Z^{-T}L^*$ qui est une matrice (n-m) x n. Par conséquent, sauf cas exeptionnel, la seule mise à jour de la réduction $G_k = Z_k^{-T}M_k Z_k^{-}$ de M_k au plan tangent $N(A_k)$, qui est d'ordre n-m, ne suffira pas pour obtenir la convergence Q-superlinéaire.

Récemment, Byrd (1985) et Yuan (1985) ont chacun donné un contre-exemple dans lequel la suite (x_k) générée par la MSQ ne converge pas Q-superlinéairement en un pas bien que le hessien réduit G^* soit parfaitement approché par la partie correspondante de M_k. Nous donnons ci-après, un autre contre-exemple qui met en évidence un comportement "gênant" de la MSQ:

$$x = [\; x^1 \quad x^2 \;]^T \;\varepsilon\; \mathbb{R}^2$$

$$f(x) = \frac{1}{2}\,||x||^2$$

$$c(x) = (x^1-1)\,(x^1+3)$$

On calcule

$$A_x = c'(x) = [\ 2(x^1+1) \quad 0\ ]$$

Si on suppose que $x^1 \neq -1$, on peut prendre

$$A_x^{-T} = [\ 1/2(x^1+1) \quad 0\ ]$$

$$Z_x = [\ 0 \quad 1\ ]$$

$$Z_x^{-T} = [\ 0 \quad 1\ ]$$

Le point optimal $x^* = [\ 1 \quad 0\ ]^T$ et $\lambda^* = -1/4$. Hessien et hessien réduit du lagrangien au point optimal s'écrivent:

$$L^* = \begin{bmatrix} 1/2 & 0 \\ 0 & 1 \end{bmatrix} \qquad G^* = 1$$

Soit $\varepsilon > 0$. On choisit

$$M_k = \begin{bmatrix} 1/2 & \varepsilon \\ \varepsilon & 1 \end{bmatrix}$$

C'est-à-dire que le hessien réduit G^* du lagrangien L^* est parfaitement approché par la matrice réduite G_k de M_k puisque l'on a $G_k = G^*$ pour tout indice k. Le déplacement est donné par (3.11) et s'écrit:

$$d_k = -\ [\ 0 \quad x_k^2\ ]^T + (x_k^1-1)(x_k^1+3)/2(x_k^1+1)\ \ [\ -1 \quad \varepsilon\ ]^T \tag{3.18}$$

$$x_{k+1} - x^* = (x_k^1-1)/2(x_k^1+1)\ \ [\ (x_k^1-1) \quad \varepsilon(x_k^1+3)\ ]^T \tag{3.19}$$

D'après le théorème de convergence local de la MSQ (voir par exemple Han (1976)), on sait qu'il existe deux constantes positives ε_0 et δ_0 telles que si $||M_k-L^*|| < \varepsilon_0$ pour tout indice k et $||x_0-x^*|| < \delta_0$, alors (x_k) converge vers x^* Q-linéairement. Choisissons ε tel que $0 < \varepsilon < (2)^{-1/2}\varepsilon_0$ et $\delta < \min(1,\delta_0)$. On a $||M_k-L^*|| \leq (2)^{1/2}\ ||M_k-L^*||_1 < \varepsilon_0$. Dès lors, en choisissant x_0 tel que $||x_0-x^*|| < \delta$, on obtiendra une suite (x_k) convergeant Q-linéairement vers x^*. Montrons que la convergence de (x_k) vers x^* n'est pas Q-superlinéaire bien que pour tout indice k on ait un choix localement optimal pour la matrice réduite G_k.

Comme x_k converge Q-linéairement vers x^*, on a $|x_k^1-1| < \delta$ et donc $1-\delta < |x_k^1| < 1+\delta$ pour tout indice k. D'après (3.19) et comme $\delta < 1$, on a pour tout indice k supérieur à 1:

$$|x_k^2| \le \varepsilon(4+\delta)/2(2-\delta)\ |x_k^1-1| \le 5\varepsilon(4-\delta)/2(2+\delta)\ |x_k^1-1| \tag{3.20}$$

Dès lors, en tenant compte de (3.19), $1-\delta < |x_k^1| < 1+\delta$ et (3.20), on obtient:

$$\begin{aligned} ||x_{k+1}-x^*|| \ge |x_k^2| &\ge \varepsilon(4-\delta)/2(2+\delta)\ |x_k^1-1| \\ &\ge \varepsilon(4-\delta)/4(2+\delta)\ |x_k^1-1| + 1/10\ |x_k^2| \\ &\ge C\ ||x_k-x^*|| \end{aligned}$$

où $C = \min(\varepsilon(4-\delta)/4(2+\delta),1/10)(2)^{-1/2}$ est une constante positive indépendante de k. Ceci montre que la convergence n'est pas Q-superlinéaire. Bien sûr, si $\varepsilon = 0$, on a $M_k = L^*$, $C = 0$ et (3.19) montre que la convergence est quadratique.

Ce contre-exemple met en évidence un effet gênant de la MSQ. En effet, supposons que x_0^2 soit nul. Dans ce cas, il suffirait de faire une suite de restaurations de la contrainte (itérations de Newton, $x_k^2 = 0$ pour tout k) pour converger quadratiquement vers x^*. Le déplacement (3.18) montre que ceci ne se produit pas avec la méthode séquentielle quadratique ($d_0^2 \neq 0$). Il y a un <u>effet déviateur</u> du fait que ε est non nul. Dans ce contre-exemple, c'est cet effet qui empêche d'obtenir la convergence Q-superlinéaire. On le supprime en prenant $\varepsilon = 0$ ou en prenant le déplacement (3.12) de la méthode de Gabay (1982,b) au lieu de (3.10). Mais comme cela apparaîtra par la suite, en général ce déplacement n'apporte pas la convergence Q-superlinéaire, même avec H_k correctement choisi.

IV - UNE METHODE A METRIQUE VARIABLE REDUITE

IV.1 - Remarques préliminaires.

De façon formelle, supposons que la matrice M_k dans la MSQ soit l'évaluation en x_k d'une application différentiable ($x \longrightarrow M(x)=M_x$). De plus, supposons que l'on connaisse exactement le hessien réduit du lagrangien au point optimal

$$G^* := Z^{-T} L^* Z^-$$

et que M_x soit telle que

$$H_x^{-1} = Z_x^{-T} M_x Z_x^- = G^*$$

pour tout x dans Ω. C'est-à-dire que la réduction de M_x au plan $N(A_x)$ est localement la meilleure possible. Cherchons alors à déterminer les modifications à apporter à la MSQ afin de remédier aux deux inconvénients suivants:

- effet déviateur du contre-exemple III.3
- absence de convergence Q-superlinéaire.

Soit φ l'application de point fixe de la MSQ, c'est-à-dire l'application telle que $x_{k+1} = \varphi(x_k)$ pour tout $k \geq 0$. (3.10) montre que l'on a

$$\varphi(x) = x - \mu_1 A_x^- c(x) - Z_x^- H g(x) + \mu_2 Z_x^- H Z_x^{-T} M_x A_x^- c(x) \tag{4.1}$$

où μ_1=1 et μ_2=1 sont des marqueurs permettant de suivre dans le calcul l'effet de leur facteur. Le facteur de μ_2 est responsable de l'effet déviateur. Pour étudier le comportement local de la MSQ, il suffit d'étudier le spectre de $T:=\varphi'(x^*)$. Par exemple, on sait que l'on aura un algorithme à convergence quadratique si T=0 (Ortega-Rheinboldt,1970). En s'aidant de la relation (2.31), on obtient:

$$\begin{aligned} T &= I - \mu_1 A^- A - Z^- H Z^{-T} L^* + \mu_2 Z^- H Z^{-T} M A^- A \\ &= (1-\mu_1) A^- A + Z^- Z - Z^- H Z^{-T} L^* Z^- Z + Z^- H Z^{-T} (\mu_2 M - L^*) A^- A \\ &= (1-\mu_1) A^- A + Z^- H Z^{-T} (\mu_2 M - L^*) A^- A \end{aligned} \tag{4.2}$$

Supposons que $\mu_1 = 1$. Alors, on montre, en considérant AT et ZT, que le spectre de T est réduit à $\{0\}$. Mais en général T n'est pas nul (il n'est pas symétrique). Par contre, on a $T^2=0$ (car $AZ^-=0$). Ceci montre la convergence quadratique en deux pas quelque soit μ_2. Remarquons également qu'il ne suffit pas de supprimer l'effet déviateur ($\mu_2=0$) pour avoir la convergence quadratique. Si M_x converge vers L^* et $\mu_1=1$, il est même essentiel de prendre $\mu_2=1$ pour avoir la convergence quadratique. L'algorithme correspondant à $\mu_1=1$ et $\mu_2=0$ a été étudié par Gabay (1982,b). Il en a montré la convergence Q-superlinéaire en deux pas.

Soient ψ l'application de point fixe d'une restauration des contraintes en x et $R:=\psi'(x^*)$. On a:

$$\psi(x) = x - A_x^- c(x) \tag{4.3}$$

$$R = I - A^-A = Z^-Z \tag{4.4}$$

Quelque soient μ_1 et μ_2 on a TR=0. Dès lors, on obtiendra un processus à convergence quadratique si on fait <u>précéder</u> le pas de la MSQ d'un pas de restauration. Et on supprimera l'effet déviateur en prenant $\mu_2=0$. Soit donc:

$$y(x) := x - A_x^- c(x) \tag{4.5}$$

$$\xi(y) := y - \mu_1 A_y^- c(y) - Z_y^- H g(y) \tag{4.6}$$

$$\chi(x) := \xi(y(x))$$

On calcule,

$$y^* := y(x^*) = x^*$$

$$\chi'(x^*) = \xi'(x^*).y'(x^*) = (I - \mu_1 A^- A - Z^- H Z^{-T} L^*) Z^- Z = 0 \tag{4.7}$$

quelque soit μ_1. Si μ_1 est nul, l'algorithme s'écrit simplement:

$$y = x - A_x^- c(x) \tag{4.8}$$

$$\chi(x) = x - A_x^- c(x) - Z_y^- H g(y) \tag{4.9}$$

En comparant (4.9) et (4.1), on voit que l'on a supprimé l'effet déviateur (μ_2=0) et que l'on a calculé le gradient réduit g en y plutôt qu'en x, c'est-à-dire après la restauration. Une itération se décompose donc en une phase de restauration des contraintes ($-A_x^- c(x)$) suivie d'une phase de minimisation du critère sur base du calcul de son gradient sur la variété $c^{-1}(c(y))$, parallèle à $c^{-1}(0)$, et munie de la métrique riemannienne:

$$\gamma_y (Z_y^- h , Z_y^- g) = (H_y^{-1} h , g)_{\mathbb{R}^{n-m}}$$

où on a supposé H_y symétrique. En effet, comme en II.4, le gradient de f sur $c^{-1}(c(y))$ pour la métrique précédente s'écrit:

$$Z_y^- H_y Z_y^{-T} f'(y) = Z_y^- H_y g(y)$$

Avec l'application χ définie en (4.9), on peut générer une suite (x_k) par $x_{k+1} = \chi(x_k)$. On génère en même temps la suite (y_k) définie à partir de (x_k) par (4.8). L'application de point fixe qui définit (y_k) s'écrit:

$$\zeta(y) = y - Z_y^- H g(y) - A_y^- c(y - Z_y^- Hg(y))$$

Au point optimal, on a $y^* = x^*$ et on obtient:

$$\zeta'(y^*) = (I - A^- A) (I - Z^- H Z^{-T} L^*) = - Z^- H Z^{-T} L^* A^- A$$

En général cet opérateur n'est pas nul, mais on a $\zeta'(y^*)^2 = 0$ (car $AZ^- = 0$). Dès lors, dans l'algorithme (4.8)-(4.9), la suite (x_k) convergera Q-linéairement et Q-superlinéairement en un pas tandis que la suite (y_k) ne convergera pas nécessairement Q-linéairement mais convergera Q-superlinéairement en deux pas. Ce résultat de convergence de la suite (y_k) a été remarqué par Coleman, Conn (1982,a) lorsque l'on prend pour A_x^- le pseudo-inverse de Penrose (2.19). Récemment, Byrd (1985) a donné un exemple dans lequel la suite (y_k) ne convergeait pas Q-superlinéairement en un pas.

IV.2 - L'algorithme modèle local.

IV.2.1 - Définition.

L'algorithme défini par (4.9) présente un inconvénient: il nécessite deux linéarisations des contraintes par itération. D'abord en x pour le calcul de A_x^- puis en y pour obtenir Z_y^-. Si on regarde le calcul (4.7), on voit que la linéarisation en y est essentielle: on ne peut pas remplacer g(y) par g(x) dans (4.9). Par contre, on peut se convaincre facilement que si on remplace A_x^- par un opérateur R_x, proche de A_x^-, ou convergeant vers A^- lorsque x converge vers x^*, on gardera la convergence Q-superlinéaire. Cette modification consiste à faire une restauration du type quasi-Newton plutôt que du type Newton.

Ces remarques conduisent à la définition de l'<u>algorithme modèle local</u>. (4.10)

1. choix de $x_0 \varepsilon \Omega$, R_0, H_0
   $k := 0$
2. linéarisation des contraintes en x_0
3. restauration des contraintes:
   $r_0 := - R_0 c_0$
   $y_0 := x_0 + r_0$
4. linéarisation des contraintes en y_0

5. minimisation:
   $q_k := - Z(y_k)^- H_k Z(y_k)^{-T} f'(y_k)$ (4.11)
   $x_{k+1} := y_k + q_k$ (4.12)
6. calcul de R_{k+1}
7. restauration des contraintes:
   $r_{k+1} := - R_{k+1} c_{k+1}$ (4.13)
   $y_{k+1} := x_{k+1} + r_{k+1}$ (4.14)
8. linéarisation des contraintes en y_{k+1}
9. mise à jour de H_k: H_{k+1}
10. $k := k+1$
    aller en 5.

On définit les déplacements:

$$d_k := r_k + q_k \tag{4.15}$$

$$h_k := q_k + r_{k+1} \tag{4.16}$$

où r_k est donné en (4.13) et q_k en (4.11). Alors les suites (x_k) et (y_k) se construisent par:

$$x_{k+1} := x_k + d_k$$

$$y_{k+1} := y_k + h_k$$

Partant d'un point x_k, on obtient le point y_k par une phase de restauration des contraintes (déplacement r_k) puis on passe au point x_{k+1} par une phase de minimisation du critère f (déplacement q_k dans le plan tangent en y_k à la variété $c^{-1}(c(y_k)))$.

IV.2.2 - <u>La phase de restauration, hypothèse (H_5).</u>

On suppose que la restauration des contraintes se fait au moyen d'un opérateur $R_k \in L(\mathbb{R}^m,\mathbb{R}^n)$ appliqué à $c_k = c(x_k)$. On fera sur (R_k) les hypothèses suivantes:

(H_5)

- (R_k) est bornée (4.17)
- AR_k est uniformément injective pour k grand: (4.18)
  il existe un indice K et une constante positive C tels que pour tout $k \geq K$ et pour tout $v \in \mathbb{R}^m$ on ait
  $||AR_k v|| \geq C \, ||v||$

(4.18) implique que R_k est uniformément injective et que $N(A) \wedge R(R_k) = \{0\}$, pour tout k plus grand que K. Comme opérateur R_k, on pourra prendre:

exemple 1: $R_k = A(x_k)^-$

exemple 2: $R_k = A(y_{k-1})^-$

exemple 3: $\eta_k R_k$ si R_k vérifie (H_5) et η_k est pris dans un compact de $(0,+\infty)$.

Si on montre qu'avec de tels R_k, la suite (x_k) converge vers x^*, alors (R_k) vérifie (H_5).

IV.2.3 - La phase de minimisation, hypothèse (H_6).

Le pas de minimisation est donné par (4.11) où les matrices H_k d'ordre n-m satisfont aux hypothèses:

$$(H_6) \quad \left| \begin{array}{ll} \cdot\ (H_k) \text{ bornée} & (4.19) \\ \cdot\ H_k \text{ non singulière} & (4.20) \end{array} \right.$$

On pose

$$G_k := H_k^{-1} \tag{4.21}$$

On ne supposera pas l'uniforme non singularité de H_k ni sa définie positivité. Remarquons que, lorsque H_k est symétrique, le pas de minimisation q_k peut s'obtenir comme solution du problème quadratique:

$$\left| \begin{array}{l} \min\limits_{q} \{ f'(y_k).q + \frac{1}{2} G_k.(Z(y_k)q)^2 \} \\ A(y_k)\ q = 0 \end{array} \right. \tag{4.22}$$

En pratique, q_k ne sera pas calculé en résolvant (4.22) mais en appliquant directement (4.11). Regardons le coût de cette opération lorsque l'on peut utiliser la technique de réduction de variables (par exemple dans les problèmes de contrôle). Dans ce cas, on prend pour $A(y_k)^-$ l'inverse à droite partitionné de $A(y_k)$ (voir (2.15)-(2.18)). Si on partitionne les vecteurs $\xi \in \mathbb{R}^n$ comme suit:

$$\xi = [\ \xi^1 \quad \xi^2\ ] \tag{4.23}$$

où $\xi^1 \in \mathbb{R}^m$ et $\xi^2 \in \mathbb{R}^{n-m}$, on calcule le multiplicateur $\lambda(y_k) = -A(y_k)^{-T} f'(y_k)$ par:

$$B(y_k)^T\ \lambda(y_k) = -\ (f'(y_k))^1 \tag{4.24}$$

Ce calcul nécessite la résolution d'un système linéaire d'ordre m. Le gradient réduit en y_k, $g(y_k) = Z(y_k)^{-T} f'(y_k)$, se calcule à partir du multiplicateur $\lambda(y_k)$:

$$g(y_k) = D(y_k)^T\ \lambda(y_k) + (f'(y_k))^2 \tag{4.25}$$

Ensuite vient la multiplication de $g(y_k)$ par $H_k = G_k^{-1}$. Enfin, comme la formule (2.18) le montre, l'application de $Z(y_k)^-$ nécessite la résolution d'un système linéaire d'ordre m:

$$B(y_k)\ q_k^1 = -\ D(y_k)\ q_k^2 \tag{4.26}$$

$$q_k^2 = H_k\ g(y_k) \tag{4.27}$$

Notons que, mise à part la multiplication par la matrice H_k, le calcul du pas de minimisation q_k correspond à la première itération d'un algorithme de gradient conjugué réduit pour résoudre (4.22). Comme ce dernier peut nécessiter (n-m) itérations (ou plus si on tient compte des imprécisions numériques), il y a un gain appréciable par rapport à la MSQ dont le déplacement se calcule en résolvant (3.2), problème dont le coût de résolution est du même ordre que celui de (4.22).

IV.3 - **Estimations.**

IV.3.1 - Proposition.

SI . c est C^2 sur Ω
. $x^* \in \Omega$ tel que $c(x^*) = 0$

ALORS pour toute suite (x_k) dans Ω on a

(i) $c_k = A\ (x_k-x^*) + O(||x_k-x^*||^2)$ (4.28)

(ii) $c(y_k) = A\ (y_k-x^*) + O(||y_k-x^*||^2)$ (4.29)

Preuve.

On utilise le développement de c autour de x^*. ●●●

IV.3.2 - Proposition.

SI . (H_0), (H_3), $x^* \in \Omega$ avec $c(x^*) = 0$
. (R_k) bornée
. (x_k) générée par l'algorithme (4.10) telle que
$x_k \in \Omega$ et $x_k \to x^*$

ALORS

(i) $r_k = - R_k A(x_k - x^*) + O(||x_k - x^*||^2)$ (4.30)

(ii) $y_k - x^* = Z^- Z(x_k - x^*) - (R_k - A^-)A(x_k - x^*) + O(||x_k - x^*||^2)$ (4.31)

(iii) $A(x_{k+1} - x^*) = - A(R_k - A^-)A(x_k - x^*) + o(||x_k - x^*||)$ (4.32)

Preuve.

(i) s'obtient à partir de (4.13) et (4.28). Comme $y_k - x^* = x_k - x^* + r_k$, (ii) se déduit de (i) en tenant compte de (2.14). D'après (ii), $y_k - x^* = O(||x_k - x^*||)$. Alors (4.11) donne en tenant compte de (2.12):

$$\begin{aligned} A(x_{k+1} - x^*) &= A(y_k)\ (x_{k+1} - x^*) + o(||x_k - x^*||) \\ &= A(y_k)\ (y_k - x^*) + o(||x_k - x^*||) \\ &= A\ (y_k - x^*) + o(||x_k - x^*||) \end{aligned}$$

d'où (iii), en utilisant (ii) et $AZ^- = 0$. ●●●

IV.3.3 - Proposition.

SI . (H_0), (H_3) et $x^* \in \Omega$ avec $g(x^*) = 0$

. (H_k) bornée

. (x_k) générée par l'algorithme (4.10) telle que $x_k \in \Omega$ et $x_k \longrightarrow x^*$

ALORS

(i) $q_k = - Z^- H_k Z^{-T} L^*(y_k - x^*) + o(||y_k - x^*||)$ (4.33)

(ii) $c_{k+1} = A(y_k - x^*) + O(||y_k - x^*||^2)$ (4.34)

(iii) $x_{k+1} - x^* = O(||y_k - x^*||)$ (4.35)

(iv) $r_{k+1} = - R_{k+1} A(y_k - x^*) + O(||y_k - x^*||^2)$ (4.36)

(v) $h_k = - (R_{k+1} A + Z^- H_k Z^{-T} L^*)(y_k - x^*) + o(||y_k - x^*||)$ (4.37)

Preuve.

En utilisant la dérivation du gradient réduit (2.31) sur (4.11) et en tenant compte de (2.24), on obtient:

$$q_k = - Z(y_k)^- H_k Z^{-T} L^* (y_k-x^*) + o(||y_k-x^*||)$$

d'où on déduit (i) puisque d'après (4.31) et la convergence de (x_k), y_k converge vers x^*. En remarquant que $c'(y_k).q_k = 0$, le développement de c_{k+1} donne:

$$c_{k+1} = c(y_k) + O(||q_k||^2)$$

On en déduit (ii) grâce à (4.29) et (i). (iii) se déduit de ce que $x_{k+1}=y_k+q_k$ et de (i). Pour (iv), on utilise la définition (4.13) de r_{k+1} et (ii). Enfin (v) s'obtient en remarquant que $h_k=q_k+r_{k+1}$ et en utilisant (i) et (iv).

●●●

IV.3.4 - Proposition.

SI . (H_0), (H_1) et (H_3)
. (H_k) bornée
. (R_k) bornée
. (x_k) générée par l'algorithme (4.10) telle que $x_k \varepsilon \Omega$ et $x_k \rightarrow x^*$

ALORS
$$h_k = Z^- Z\ d_k - (R_k - A^-) A\ d_k + (R_{k+1} - R_k) A (R_k - A^-) A (x_k - x^*) + o(||x_k-x^*||) \qquad (4.38)$$

Preuve.
D'après (4.13) et (4.28), on a

$$y_k = x_k - R_k A (x_k - x^*) + O(||x_k-x^*||^2) \qquad (4.39)$$

Comme (4.35) et (4.31) montrent que $x_{k+1}-x^* = O(||x_k-x^*||)$, on a en faisant la différence de (4.39) prise avec l'indice k+1 et l'indice k:

$$h_k = d_k - R_k A d_k - (R_{k+1} - R_k) A (x_{k+1} - x^*) + O(||x_k-x^*||^2)$$

On conclut en utilisant (4.32).

●●●

Les équivalences suivantes demandent sur la suite (H_k) une hypothèse un peu plus forte que (4.20) à savoir son uniforme non singularité.

IV.3.5 - Proposition.

SI . (H_0), (H_1), (H_2) et (H_3)
. (x_k) générée par l'algorithme (4.10) avec (H_5), (H_6) telle que $x_k \in \Omega$ et $x_k \longrightarrow x^*$
. (H_k) uniformément non singulière
ALORS 1) $\|\|d_k\|\| \sim \|\|x_k - x^*\|\|$ (4.40)
2) $\|\|h_k\|\| \sim \|\|y_k - x^*\|\|$ (4.41)

Preuve.

1) Comme $d_k = r_k + q_k$, (4.30), (4.33) et (4.31) donnent:

$$d_k = - [R_kA + Z^- H_k Z^{-T} L^*(I-R_kA)]\ (x_k-x^*) + o(||x_k-x^*||) \tag{4.42}$$

(4.40) se déduira alors de l'uniforme non singularité de l'opérateur entre crochets qui est d'ordre n. Pour montrer cela, il suffit de montrer qu'il existe une constante positive C et un entier K_1 tels que pour tout $\xi \in \mathbb{R}^n$ et pour tout indice k supérieur à K_1, on ait:

$$||\ [\ R_kA + Z^- H_k Z^{-T} L^*\ (I-R_kA)\ ]\ \xi\ || \geq C\ ||\ \xi\ || \tag{4.43}$$

Soit K donné par l'hypothèse (H_5). Si (4.43) n'est pas vérifiée, il existe une sous-suite d'indices $\{k_i : i \geq 0\}$ et une suite $\{\xi_i : i \geq 0\}$ dans $\mathbb{R}^n$, telles que:

$$k_i \geq K \quad \text{pour tout } i \geq 0$$

$$||\ \xi_i\ || = 1 \quad \text{pour tout } i \geq 0$$

$$[\ R_{k_i} A + Z^-\ H_{k_i}\ Z^{-T}\ L^*\ (I - R_{k_i} A)\ ]\ \xi_i \longrightarrow 0 \tag{4.44}$$

lorsque i tend vers l'infini. En multipliant (4.44) par A, on obtient:

$$A\ R_{k_i}\ A\ \xi_i \longrightarrow 0$$

D'après (H_5), $\overline{AR_k}$ est uniformément injective pour $k \geq K$. Dès lors,

$$A\ \xi_i \longrightarrow 0 \qquad (4.45)$$

Comme $\xi_i = A^- A\ \xi_i + Z^- Z\ \xi_i$, (4.44) et (4.45) donnent:

$$Z^-\ H_{k_i}\ Z^{-T}\ L^*\ Z^-\ Z\ \xi_i \longrightarrow 0$$

L'injectivité de Z^-, l'uniforme injectivité de H_k et la non singularité de G^* implique que

$$Z\ \xi_i \longrightarrow 0$$

Ce qui joint à (4.45) montre que ξ_i converge vers 0, en contradiction avec $||\xi_i||=1$.

2) De la même manière, on obtient (4.41) à partir de (4.37) et en montrant l'uniforme non singularité de

$$R_{k+1}\ A + Z\ H_k\ Z^{-T}\ L^*$$

pour k assez grand. ●●●

IV.4 - Convergence locale.

Le théorème suivant étudie un pas de l'algorithme (4.10) lorsque (x_k, R_k, G_k) est choisi proche de (x^*, A^-, G^*).

IV.4.1 - Théorème.

SI . (H_0), (H_1), (H_2) et (H_3)

. r un réel positif

ALORS il existe des réels positifs ε, β et δ tels que si $x_k \in \Omega$, R_k et G_k vérifient

$$||x_k - x^*|| < \varepsilon \qquad (4.46)$$

$$||R_k - A^-|| < \beta \qquad (4.47)$$

$$||G_k - G^*|| < \delta \qquad (4.48)$$

on a

(i) G est non singulière et $H_k = G_k^{-1}$ est majorée en norme par une constante $C=C(G^*,\delta)$ indépendante de r, ε et β. (ii) $\|x_{k+1}-x^*\| \leq r\ \|x_k-x^*\|$ où x_{k+1} est calculé à partir de x_k par l'algorithme (4.10).

Preuve.

1) Soit $r > 0$. On fixe ε_0 et β_0 deux constantes positives. On choisira $\varepsilon \leq \varepsilon_0$ et $\beta \leq \beta_0$. Soit $\gamma_0 > 0$ tel que

$$\| (G^*)^{-1} \| \leq \gamma_0 \tag{4.49}$$

On fixe $\delta_0 < 1/\gamma_0$ et on choisira $\delta \leq \delta_0$. Dès lors, (4.49) et $\|G_k - G^*\| < \delta \leq 1/\gamma_0$ impliquent que G_k est inversible et $H_k = G_k^{-1}$ vérifie:

$$\| H_k \| \leq 1 / (1/\gamma_0 - \delta) \tag{4.5$_0$}$$

Ceci montre (i).

2) On désigne par C une constante positive "générique" dépendant de ε_0, β_0 et δ_0 et indépendante de r, ε, β et δ. ε, β et δ seront ajustés ultérieurement. De l'identité

$$c_k = A\ (x_k - x^*) + \int_0^1 [\ c'(x^* + t(x_k - x^*)) - c'(x^*)\ ].(x_k - x^*)\ dt$$

on déduit,

$$\|c_k\| \leq C\ \|x_k - x^*\| \tag{4.51}$$

$$\|c_k - A\ (x_k - x^*)\ \| \leq C\ \|x_k - x^*\|^2 \tag{4.52}$$

$$\|y_k - x_k\| = \|r_k\| \leq C\ \|x_k - x^*\| \tag{4.53}$$

$$\|y_k - x^*\| \leq C\ \|x_k - x^*\| \tag{4.54}$$

3) Grâce à l'identité

$$I = A(y_k)^-\ A(y_k) + Z(y_k)^-\ H_k\ G_k\ Z(y_k)$$

et à l'expression du déplacement d_k, on obtient:

$$x_{k+1} - x^* = [\ A(y_k)^- A(y_k)(x_k - x^*) - R_k c_k] \tag{4.55}$$

$$+ Z(y_k)^- H_k \; [\; G_k Z(y_k)(x_k - x^*) - g(y_k) \;]$$

. Le premier terme du second membre de (4.55) s'écrit:

$$A(y_k)^- \; [\; A(y_k) - A \;]\; (x_k - x^*) + A(y_k)^- \; [\; A\,(x_k - x^*) - c_k \;]$$
$$+ [\; A(y_k)^- - A^- \;]\; c_k + [\; A^- - R_k \;]\; c_k$$

Grâce aux estimations (4.54), (4.52) et (4.51), ce terme est majoré par

$$C\; ||x_k - x^*||^2 + C\,\beta\; ||x_k - x^*|| \tag{4.56}$$

D'autre part, en développant $g(y_k)$ autour de x^* à l'aide de la relation (2.31) et en tenant compte de ce que $g(x^*)=0$, on obtient:

$$\begin{aligned} & ||\; G_k\; Z(y_k)\; (x_k - x^*) - g(y_k)\; || \\ \le\; & ||\; G_k\; Z(y_k)\; (x_k - x^*) - Z^{-T}\, L^*\; (y_k - x^*)\; || + \eta(\varepsilon)\; ||y_k - x^*|| \\ \le\; & ||\; G_k\; Z\; (x_k - x^*) - Z^{-T}\, L^*\; (x_k - x^*) + Z^{-T}\, L^*\; R_k\; A\; (x_k - x^*)\; || \\ & \; + \eta(\varepsilon)\; ||x_k - x^*|| + C\; ||x_k - x^*||^2 \end{aligned} \tag{4.57}$$

où, grâce à (4.54), $\eta(\varepsilon)$ converge vers 0 lorsque ε tend vers 0. Avec (4.56) et (4.57), (4.55) devient

$$||x_{k+1} - x^*|| \le C\; (\varepsilon + \eta(\varepsilon) + \beta + \delta\;)\; ||x_k - x^*||$$

C'est-à-dire (ii) en choisissant ε, β et δ assez petit. ●●●

Le corollaire suivant donne la convergence Q-linéaire locale de l'algorithme (4.10).

IV.4.2 - <u>Corollaire.</u>

SI . (H_0), (H_1), (H_2) et (H_3)

. $r \in\;]0,1[$

ALORS il existe des réels positifs ε, β et δ tels que

si $x_0 \in \Omega$, R_k, G_k vérifient

$$||x_0 - x^*|| < \varepsilon \tag{4.58}$$

$$||R_k - A^-|| < \beta, \text{ pour tout indice } k \tag{4.59}$$

$$||G_k - G^*|| < \delta, \text{ pour tout indice } k \tag{4.60}$$

on a

(i) G_k est non singulière pour tout indice k

(ii) (G_k) et (H_k) sont bornées, $(H_k := G_k^{-1})$

(iii) la suite (x_k) générée par l'algorithme (4.10) est bien définie et converge Q-linéairement vers x^*: pour tout indice k, on a

$$||x_{k+1}-x^*|| \leq r \; ||x_k - x^*||$$

Remarque.

On ne suppose pas les hypothèses (H_5) et (H_6) sur R_k et H_k, mais comme le montrent (i) et (ii), une partie de (H_5) est contenue dans (4.59) et (H_6) est contenu dans (4.60).

Preuve.

Soit $r > 0$. Soient ε_1, β, δ donnés par le théorème IV.4.1. On choisit $\varepsilon \leq \varepsilon_1$ tel que $B(x^*,\varepsilon) \subset \Omega$. Le résultat s'obtient par récurrence en utilisant le théorème IV.4.1. ●●●

Le corollaire suivant donne une condition suffisante de convergence Q-superlinéaire.

IV.4.3 - Corollaire.

SI . (H_0), (H_1), (H_2) et (H_3)

. $R_k \longrightarrow A^-$

. $G_k \longrightarrow G^*$

ALORS il existe des réels positifs ε, β et δ tels que si $x_0 \in \Omega$, R_k, G_k vérifient (4.58), (4.59) et (4.60) alors la suite (x_k) générée par l'algorithme (4.10) est bien définie et converge Q-superlinéairement vers x^*.

Preuve.

D'après le corollaire IV.4.2, la suite (x_k) est bien définie et converge Q-linéairement vers x^* si on choisit ε, β et δ assez petits. Soit $r > 0$ et ε_r, β_r et δ_r donnés par le corollaire IV.4.2. De la convergence des suites (x_k), (R_k) et (G_k), on déduit l'existence d'un indice K_r tel que:

$$||x_{K_r} - x^*|| < \varepsilon_r$$

$$||R_k - A^-|| < \beta_r, \text{ pour tout indice } k \geq K_r$$

$$||G_k - G^*|| < \delta_r, \text{ pour tout indice } k \geq K_r$$

Dès lors, d'après le corollaire IV.4.2, on a

$$||x_{k+1} - x^*|| \leq r \; ||x_k - x^*||$$

pour tout indice $k \geq K_r$. Comme ceci est vrai pour tout choix de r, le corollaire est démontré. ●●●

IV.5 - CNS de convergence Q-superlinéaire.

IV.5.1 - Théorème.

SI . on a (H_0), (H_1), (H_2) et (H_3)

. (x_k) générée par l'algorithme (4.10) avec (H_5) et (H_6) telle que $x_k \in \Omega$ et $x_k \to x^*$.

. $ZR_k \to 0$ (4.61)

ALORS les clauses suivantes sont équivalentes

(i) la convergence de x_k vers x^* est Q-superlinéaire

(ii) $A^-A\,(x_k - x^*) + r_k = o(||x_k - x^*||)$ (4.62)

$Z^-Z\,(x_k - x^*) + q_k = o(||x_k - x^*||)$ (4.63)

(iii) $(R_k - A^-)\,A\,(x_k - x^*) = o(||x_k - x^*||)$ (4.64)

$(G_k - G^*)\,Z\,(x_k - x^*) = o(||x_k - x^*||)$ (4.65)

(iv) $(R_k - A^-)\,A\,d_k = o(||x_k - x^*||)$ (4.66)

$(G_k - G^*)\,Z\,d_k = o(||x_k - x^*||)$ (4.67)

(v) $c_{k+1} = o(||x_k - x^*||)$ (4.68)

$$\boxed{g(y_{k+1}) = o(||x_k - x^*||)} \tag{4.69}$$

Remarques.

Le théorème ne traite que de l'itération k et mis à part (4.69) ne fait intervenir que des grandeurs calculées pour passer de x_k à x_{k+1}. Dès lors le résultat se généralise aisément à des sous-suites de (x_k). Par exemple, si (k_i) est une sous-suite d'indices, on a équivalence entre

$$\frac{||x_{k_i+1} - x^*||}{||x_{k_i} - x^*||} \longrightarrow 0 \qquad \text{quand } i \longrightarrow \infty$$

et les deux clauses suivantes:

$$\frac{||(R_{k_i} - A^-) A (x_{k_i} - x^*)||}{||x_{k_i} - x^*||} \longrightarrow 0 \qquad \text{quand } i \longrightarrow \infty$$

$$\frac{||(G_{k_i} - G^*) Z (x_{k_i} - x^*)||}{||x_{k_i} - x^*||} \longrightarrow 0 \qquad \text{quand } i \longrightarrow \infty$$

Si en plus de (H_6), on suppose l'uniforme non singularité des H_k alors l'équivalence (4.40) entre $||d_k||$ et $||x_k - x^*||$ montre que l'on peut remplacer $o(||x_k - x^*||)$ par $o(||d_k||)$ dans (4.62) à (4.69).

On utilisera dans la démonstration, l'estimation

$$|| d_k || \leq C || x_k - x^* ||$$

qui peut s'obtenir à partir de (4.42).

Preuve.

1) En utilisant (4.61), on voit que l'on a:

$$Z (R_k - A^-) A (x_k - x^*) = Z R_k A (x_k - x^*) = o(||x_k - x^*||) \tag{4.70}$$

2) On montre les équivalences: (4.62) <=> (4.64) <=> (4.66) <=> (4.68).

- (4.62) est équivalente à (4.64) car d'après (4.30) on a

$$r_k + A^- A\ (x_k - x^*) = -\ (R_k - A^-)\ A\ (x_k - x^*) + o(||x_k - x^*||) \tag{4.71}$$

- D'après (4.42),

$$\begin{aligned} A\ d_k &= -\ A\ R_k\ A\ (x_k - x^*) + o\ (||x_k - x^*||) \\ (R_k - A^-)\ A\ d_k &= -\ (R_k A R_k A - A^- A R_k A)\ (x_k - x^*) + o(||x_k - x^*||) \end{aligned} \tag{4.72}$$

mais, grâce à (4.61) on a

$$\begin{aligned} A^-\ A\ R_k\ A\ (x_k - x^*) &= (A^- A + Z^- Z)\ R_k\ A\ (x_k - x^*) + o(||x_k - x^*||) \\ &= R_k\ A\ A^-\ A\ (x_k - x^*) + o(||x_k - x^*||) \end{aligned}$$

puisque $AA^- = I$. (4.72) devient:

$$(R_k - A^-)\ A\ d_k = -\ R_k\ A\ (R_k - A^-)\ A\ (x_k - x^*) + o(||x_k - x^*||)$$

Comme (R_k) est bornée et uniformément injective, on a l'équivalence

$$(4.66) \iff A\ (R_k - A^-)\ A\ (x_k - x^*) = o(||x_k - x^*||) \tag{4.73}$$

On obtient alors l'équivalence entre (4.64) et (4.66) en utilisant (4.70).

- D'après (4.34) et (4.31), on a

$$\begin{aligned} c_{k+1} &= -\ A\ (R_k - A^-)\ A\ (x_k - x^*) + o(||x_k - x^*||) \\ &= -\ (R_k - A^-)\ A\ (x_k - x^*) + o(||x_k - x^*||) \end{aligned} \tag{4.74}$$

où on a utilisé (4.70). Cette dernière relation montre l'équivalence entre (4.74) et (4.67).

3) (ii) <= (iii). En utilisant (4.33) et (4.31) et puis (4.64), on obtient:

$$q_k + Z^-\ Z\ (x_k - x^*) = Z^-\ H_k\ (G_k - G^*)\ Z\ (x_k - x^*)$$

$$+ Z^- H_k Z^{-T} L^* (R_k-A^-) A (x_k-x^*) + o(||x_k-x^*||) \qquad (4.75)$$
$$= Z^- H_k (G_k-G^*) Z.(x_k-x^*) + o(||x_k-x^*||)$$

On obtient alors (ii), grâce à (4.65).

4)<u>(i) <=> (ii)</u>. Comme $x_{k+1} = x_k + r_k + q_k$ et $A^-A + Z^-Z = I$, on obtient:

$$x_{k+1}-x^* = r_k + A^-A (x_k-x^*) + q_k + Z^-Z (x_k-x^*) \qquad (4.76)$$

(4.76) montre l'implication (ii) => (i). Si on suppose que l'on a (i), en multipliant (4.76) par Z^-Z et en tenant compte de (4.71) et (4.70), on obtient:

$$Z^-Z [q_k + Z^-Z (x_k-x^*)] = o (||x_k-x^*||)$$

Ce qui, d'après (4.75), implique (4.63). Alors (i) et (4.63) injectés dans (4.76) implique (4.62).

5)<u>(iv) <=> (v)</u>. D'après l'expression du déplacement, on obtient:

$$- (G_k-G^*) Z d_k = - G_k Z(y_k) d_k + G^* Z d_k + o(||x_k-x^*||)$$
$$= G_k Z(y_k) R_k c_k + g(y_k) + G^* Z d_k + o(||x_k-x^*||)$$

Avec (4.61), on obtient:

$$\begin{aligned} - (G_k-G^*) Z d_k &= g(y_k) + G^* Z d_k + o(||x_k-x^*||) \\ &= g(y_{k+1}) - \int_0^1 g'(y_k+th_k).h_k \, dt + G^* Z d_k + o(||x_k-x^*||) \\ &= g(y_{k+1}) + Z^{-T} L^* (Z^-Z d_k - h_k) + o(||x_k-x^*||) \\ &= g(y_{k+1}) + Z^{-T} L^* (R_k-A^-) A d_k \\ &\quad - Z^{-T} L^* (R_{k+1}-R_k) A (R_k-A^-) A (x_k-x^*) + o(||x_k-x^*||) \qquad (4.77) \end{aligned}$$

où on a utilisé (4.38). Si (iv) ou (v) est vérifiée alors selon la deuxième étape, (4.64) et (4.66) sont vérifiées. (4.77) s'écrit alors:

$$- (G_k-G^*) Z d_k = g(y_{k+1}) + o(||x_k-x^*||)$$

relation qui montre l'équivalence entre (iv) et (v).

6)<u>(v) <=> (i)</u>. Comme (i) est équivalente à (ii), si (i) ou (v) est vérifiée, on a (4.64). Dès lors, (4.32) montre que l'on a

$$A\ (x_{k+1} - x^*) = o(||x_k - x^*||) \tag{4.78}$$

En utilisant (2.31), (4.31) et (4.78), on obtient:

$$\begin{aligned} g(y_{k+1}) &= Z^{-T}\ L^*\ (y_{k+1} - x^*) + o(||x_k - x^*||) \\ &= Z^{-T}\ L^*\ Z^-\ Z\ (x_{k+1} - x^*) + o(||x_k - x^*||) \end{aligned} \tag{4.79}$$

Si (i) est vérifiée, on a (4.68) grâce aux étapes 4 et 2, et on a (4.69) grâce à (4.79). Inversément, si (v) est vérifiée, (4.79) et la non singularité de $Z^{-T}L^*Z^-$ montre que l'on a:

$$Z\ (x_{k+1} - x^*) = o(||x_k - x^*||)$$

ce qui, joint à (4.78) donne la convergence Q-superlinéaire de (x_k).

7)<u>(iii) <=> (iv)</u>. D'après les implications précédentes, si (iii) ou (iv) est vérifiée, la suite (x_k) converge Q-superlinéairement vers x^*. Alors, on a

$$d_k = -\ (x_k - x^*) + o(||x_k - x^*||)$$

Cette équation suffit pour montrer l'équivalence entre (iii) et (iv).

8)<u>Conclusion</u>. On a montré les implications (i) <=> (ii) <= (iii) <=> (iv) <=> (v) <=> (i). Ce qui suffit. ●●●

V - GLOBALISATION DE LA METHODE

V.1 - Hypothèse supplémentaire.

Rappelons que les hypothèses (H_0), (H_1), (H_2) et (H_3) ont été introduites à la section II et les hypothèses (H_5) et (H_6) à la section IV. On supposera parfois plus de régularité sur f et c par l'hypothèse suivante:

(H_0') | On suppose que (H_0) est vérifiée et que f et c sont de classe C^3 sur Ω

Cette régularité permet de choisir A_x, A_x^-, Z_x, Z_x^- fonctions C^2 de x sur Ω (sous l'hypothèse (H_3)). En particulier, l'application

$$x \in \Omega \longrightarrow g'(x) := (Z_x^{-T} f'(x))'$$

sera lipschitzienne.

On supposera dans cette section que $\Omega = \mathbb{R}^m$.

V.2 - Obtention de l'algorithme modèle global.

V.2.1 - Introduction.

Comme on le verra à la section VI, il est possible de construire des algorithmes à partir de l'algorithme modèle (4.10) ayant un bon comportement local: si on choisit le couple (x_0, G_0) suffisamment proche de (x^*, G^*), la suite générée (x_k) convergera Q-superlinéairement vers x^*. En pratique, (x_0, G_0) n'est pas nécessairement proche de (x^*, G^*) et le comportement des algorithmes n'est pas prévisible. Il est donc important pour la viabilité de la méthode de montrer qu'il est possible de la "globaliser", c'est-à-dire de rendre la convergence des suites générées indépendante du choix du couple initial (x_0, G_0) et cela sous des conditions raisonnables. C'est ce dont nous nous occuperons dans cette section.

Un autre point important est le raccord entre la méthode globale (celle assurant la convergence globale) et la méthode locale (celle de la section IV). Pour que la convergence Q-superlinéaire soit assurée il faut, lorsque x_k est proche de x^*, que le déplacement calculé par la méthode globale soit le déplacement d_k de l'algorithme (4.10). On sait que ceci n'est pas vérifié dans la globalisation de la MSQ proposée par Han (1977): c'est ce qu'on appelle l'effet Maratos (Maratos (1978), Chamberlain, Powell, Lemarechal, Pedersen (1982)). Dans la globalisation présentée ci-après, cet effet n'apparaît pas à condition que G^* soit, dans un certain sens, bien approchée par G_k. Cette propriété dépend donc de la méthode de mise à jour utilisée.

V.2.2 - Obtention de l'algorithme.

La méthode proposée par Han (1977) pour globaliser la MSQ est basée sur la pénalisation exacte du critère f. Soit p un réel positif. On note θ_p la fonctionnelle pénalisée exacte de f définie par:

$$\theta_p(x) = f(x) + p\,||c(x)||_1 \tag{5.1}$$

où $||.||_1$ désigne la 1-norme de $\mathbb{R}^m$. On dit que la pénalisation est exacte parce que si le paramètre de pénalisation p est pris assez grand (supérieur à $||\lambda^*||_\infty$), tout point stationnaire de (1.1) est point stationnaire de θ_p et tout point vérifiant les conditions suffisantes d'optimalité du second ordre de (1.1) est minimum local strict de θ_p (Fletcher (1981)). Dans ce dernier cas, on peut chercher x^* en minimisant θ_p. Pour la MSQ, le point essentiel est de constater que si p est assez grand et M_k définie positive, le déplacement d_k calculé en x_k par la MSQ (c.f.r. (3.10)) est une direction de descente de θ_p en x_k. On peut donc obtenir les directions de descente successives de cette manière sans avoir à dériver θ_p qui n'est pas différentiable. A chaque itération, on fait ensuite sur θ_p une recherche linéaire en x_k dans la direction d_k. Ceci détermine le pas ρ_k et l'itéré suivant $x_{k+1} = x_k + \rho_k d_k$. Cette façon de procéder a un autre avantage: si le pas déterminé par la recherche linéaire s'avére asymptotiquement être le pas unité, alors $x_{k+1} = x_k + d_k$ et le déplacement étant donné par d_k, la convergence a quelque chance d'être Q-superlinéaire (il faut également que des conditions d'approximation de L^* par M_k soient vérifiées).

Dans notre cas, la globalisation s'obtiendra également en cherchant à minimiser Θ_p, mais il y a une difficulté supplémentaire. En effet, on peut décider si une direction est de descente pour Θ_p en un point x si elle n'utilise dans sa définition que des grandeurs calculées en x: c'est une propriété locale en x. Par exemple, le déplacement $d_k = r_k + q_k$ donné par l'algorithme modèle (4.10) est une direction issue de x_k, mais comme q_k donné par (4.11) est défini avec des grandeurs calculées en y_k, il n'est pas possible de dire si, en x_k, Θ_p décroît dans la direction d_k. De même pour $h_k = q_k + r_{k+1}$ qui est issue de y_k. Cette fois c'est r_{k+1} donné par (4.13) qui fait intervenir le point x_{k+1} plutôt que le point y_k. On ne peut donc pas non plus dire si h_k est une direction de descente de Θ_p en y_k. En ce qui concerne r_k, cela dépend du choix de l'opérateur R_k. Si $R_k = A(x_k)^-$, $r_k \neq 0$ et $p > ||\lambda(x_k)||_\infty$ alors on peut montrer que r_k est une direction de descente de Θ_p en x_k (c.f.r. les preuves ci-après). Tandis que si $R_k = A(y_{k-1})^-$, on ne peut pas se prononcer. Enfin q_k est une direction de descente de Θ_p en y_k si et seulement si q_k est non nul. Si q_k est nul, $x_{k+1} = y_k$ et $r_{k+1} = -A^-_{k+1}\ c_{k+1}$ ou $r_{k+1} = -A(y_k)^-\ c_{k+1}$ sont identiques; si de plus $p > ||\lambda(y_k)||_\infty$ alors r_{k+1} est une direction de descente de Θ_p en y_k. Ces considérations nous conduisent à définir en y_k un <u>arc de descente</u> de Θ_p, tangent à q_k en y_k:

$$y_k(\rho) := y_k + \rho\, q_k + \rho^2\, r_{k+1}, \quad 0 < \rho \leq 1 \tag{5.2}$$

où q_k et r_{k+1} sont définis comme précédemment:

$$q_k := -\, Z(y_k)^-\ H_k\ Z(y_k)^{-T}\ f'(y_k) \tag{5.3}$$

$$r_{k+1} := -\, R_{k+1}\ c(x_{k+1}) \tag{5.4}$$

On définit y_{k+1} à partir de y_k par un choix particulier de ρ:

$$y_{k+1} := y_k(\rho_k) \tag{5.5}$$

$$h_k := q_k + r_{k+1} \tag{5.6}$$

$$v_k := y_{k+1} - y_k \tag{5.7}$$

On définit également x_{k+1} à partir de y_k:

$$x_{k+1} := y_k + q_k \tag{5.8}$$

$$d_k := r_k + q_k \tag{5.9}$$

$$u_k := x_{k+1} - x_k \tag{5.10}$$

Lorsque ρ_{k-1} est différent de 1, $d_k \neq u_k$ et on ne peut plus définir x_{k+1} à partir du seul point x_k. L'étude de la suite (x_k) n'est donc pas aisée et dans cette section, on s'attachera plutôt à la suite (y_k). Dans (5.5), ρ_k est déterminé de telle sorte que le passage de y_k à y_{k+1} fasse décroître Θ_p. Pour cela, on utilise une règle du type Armijo (1966):

$$\beta \in]0,1[\tag{5.11}$$

$$\rho_k := \beta^{l_k} \tag{5.12}$$

où l_k est le plus petit entier naturel tel que

$$\begin{aligned} \Theta_p(y_k(\beta^{l_k})) < \Theta_p(y_k) &- \beta^{l_k} \alpha (H_k g(y_k) , g(y_k)) \\ &- \beta^{2l_k} \alpha (p - ||\lambda(y_k)||_\infty) ||c(y_k)||_1 \end{aligned} \tag{5.13}$$

Dans (5.13), α est un réel qui sera choisi dans $]0,1/2[$ pour des raisons qui apparaîtront à la fin de cette section. La règle d'Armijo (5.11)-(5.12) a l'avantage de donner à ρ_k la valeur 1 si cela est possible pour le critère (5.13) (la première valeur essayée dans (5.13) est $l_k=0$, $\rho_k=1$) et donc de ne pas empêcher une convergence Q-superlinéaire éventuelle pour la suite (x_k).

L'objet du théorème suivant est de donner des conditions d'une part pour que l'arc défini par (5.2) soit bien un arc de descente de Θ_p en y_k et d'autre part pour qu'il existe un entier naturel l_k tel que (5.13) soit vérifiée. On note:

$$\lambda(x) := - A(x)^{-T} f'(x) \tag{5.14}$$

V.2.3 - <u>Théorème.</u>

SI . (H_0)

. $0 < \alpha < 1$

. $p > ||\lambda(y_k)||_\infty$ (5.15)

. H_k définie positive

. $R_{k+1} = A(y_k)^-$ si $x_{k+1} = y_k$ (5.16)

ALORS y_k n'est pas un point stationnaire de (1.1) SSI il existe un réel positif $\bar{\rho}$ tel que pour tout ρ dans $]0,\bar{\rho}]$ on ait:

$$\begin{aligned}\Theta_p(y_k(\rho)) < \Theta_p(y_k) &- \rho\ \alpha\ (H_k g(y_k), g(y_k)) \\ &- \rho^2\ \alpha\ (p - \|\lambda(y_k)\|_\infty)\ \|c(y_k)\|_1\end{aligned} \tag{5.17}$$

<u>Preuve.</u>

Si y_k est un point stationnaire de (1.1), $c(y_k) = 0$ et $g(y_k) = 0$. Dès lors, on a $q_k = 0$, $r_{k+1} = 0$ et $y_k(\rho) = y_k$ quelque soit ρ. Ceci montre que l'on a égalité dans (5.17). Inversément, supposons qu'il existe une suite (ρ_i) de réels positifs tels que $\rho_i \longrightarrow 0$ et que pour tout indice i on ait $0 < \rho_i \leq 1$ et

$$\begin{aligned}\Theta_p(z_i) \geq \Theta_p(y_k) &- \rho_i\ \alpha\ (H_k g(y_k), g(y_k)) \\ &- \rho_i^2\ \alpha\ (p - \|\lambda(y_k)\|_\infty)\ \|c(y_k)\|_1\end{aligned} \tag{5.18}$$

où on a posé $z_i := y_k(\rho_i)$. On a

$$\begin{aligned}\Theta_p(z_i) &= f(y_k) + \rho_i\ f'(y_k).q_k + O(\rho_i^2) + p\ \|c(y_k) + O(\rho_i^2)\|_1 \\ &\leq \Theta_p(y_k) - \rho_i\ (H_k g(y_k), g(y_k)) + O(\rho_i^2)\end{aligned} \tag{5.19}$$

où on a utilisé le fait que $c'(y_k).q_k = 0$ puis (4.11) et (2.29). Si on injecte (5.19) dans (5.18) et qu'on passe à la limite en i après avoir divisé par ρ_i, on trouve:

$$(1-\alpha)\ (H_k g(y_k), g(y_k)) \leq 0$$

Comme $(1-\alpha) > 0$ et H_k est définie positive, cela implique successivement que

$$g(y_k) = 0, \quad q_k = 0, \quad x_{k+1} = y_k \tag{5.20}$$

D'après (5.16) et (5.20), $r_{k+1} = - A(y_k)^{-}c(y_k)$. On développe à nouveau $\Theta_p(z_i)$ en tenant compte de (5.20):

$$\begin{aligned}\Theta_p(z_i) &= f(y_k) + \rho_i^2\ f'(y_k).r_{k+1} + O(\rho_i^4) \\ &\quad + p\ \|\ c(y_k) + \rho_i^2\ c'(y_k).r_{k+1} + O(\rho_i^4)\ \|_1 \\ &= f(y_k) + \rho_i^2\ (\lambda(y_k), c(y_k)) + O(\rho_i^4) \\ &\quad + p\ \|\ (1-\rho_i^2)\ c(y_k) + O(\rho_i^4)\ \|_1\end{aligned}$$

$$\leq \theta_p(y_k) - \rho_i^2 \ (p - ||\lambda(y_k)||_\infty) \ ||c(y_k)||_1 + O(\rho_i^4) \tag{5.21}$$

On a également utilisé le fait que $0 < \rho_i \leq 1$. Si on injecte (5.21) dans (5.18) en tenant compte de (5.20) et qu'on passe à la limite en i après avoir divisé par ρ_i^2, on obtient:

$$(1-\alpha) \ (p - ||\lambda(y_k)||_\infty) \ ||c(y_k)||_1 \leq 0 \tag{5.22}$$

Comme $(1-\alpha) > 0$ et $(p-||\lambda(y_k)||_\infty) > 0$, (5.22) montre que $c(y_k)=0$. Ce qui joint à (5.20) montre que y_k est point stationnaire de (1.1). ●●●

V.2.4 - <u>Choix du paramètre de pénalisation.</u>

L'inégalité (5.15) montre qu'il faudra adapter p à certaines itérations. On notera p_k le paramètre de pénalisation utilisé à l'itération k. On supposera dans la suite que la règle d'adaptation de p_k satisfait aux trois conditions suivantes:

1) $p_k \geq ||\lambda(y_k)||_\infty + \bar{p}$, pour tout indice k (5.23)

2) il existe un indice K' tel que pour tout indice $k \geq K'$,
$(p_{k-1} \geq ||\lambda(y_k)||_\infty + \bar{p}) \Rightarrow (p_k = p_{k-1})$ (5.24)

3) (p_k) est bornée SSI p_k est modifié un nombre fini de fois (5.25)

Dans (5.23) et (5.24), $\bar{p}$ est une constante positive fixée. (5.24) signifie qu'asymptotiquement $(k \geq K')$, on ne modifie p_{k-1} que si (5.23) n'est pas vérifiée. Il s'en suit qu'au delà de K', (p_k) est croissante. Par exemple, on pourra prendre le schéma suivant, inspiré de Mayne, Polak (1982):

<u>si</u> $p_{k-1} \geq ||\lambda(y_k)||_\infty + \bar{p}$ <u>alors</u> $p_k := p_{k-1}$ (5.26)

<u>sinon</u> $p_k := \max (\delta \, p_{k-1} \, , \, ||\lambda(y_k)||_\infty + \bar{p})$ (5.27)

Dans (5.27), δ est une constante fixée strictement supérieure à 1, afin de satisfaire à la condition (5.25).

L'hypothèse (5.16) sera vérifiée si on prend $R_k = A(y_{k-1})^-$ ou $R_k = A(x_k)^-$ et c'est ce qui sera fait dans les algorithmes concrets de la section VI. Avec ces opérateurs R_k, on obtient une propriété que n'impliquent pas ceux qui vérifient (H_5) et qui est donnée dans le lemme suivant.

V.2.5 - Lemme.

SI R_k converge vers un opérateur injectif R	(5.28)
ALORS $\|\|r_k\|\| \sim \|\|c(x_k)\|\|$ pour k grand	

Preuve.

D'après (5.28), (R_k) est bornée. Comme $r_k = - R_k\, c(x_k)$, on a $r_k = O(||c(x_k)||)$. Inversément, (5.28) montre que l'on a

$$r_k = - R\, c(x_k) + o(||c(x_k)||)$$

Ce qui, avec l'injectivité de R, montre que $c(x_k) = O(||r_k||)$ pour k grand. ●●●

V.2.6 - L'algorithme global.

On peut à présent définir l'algorithme modèle global: (5.29)

1. choix de $x_0 \in \Omega$ et H_0 définie positive
   $k := 0$
2. linéarisation des contraintes en x_0
3. restauration des contraintes:
   $y_0 := x_0 - A(x_0)^- \, c(x_0)$
4. linéarisation des contraintes en y_0:
   $\lambda(y_0)$ par (5.14)
5. minimisation:
   q_k par (5.3), x_{k+1} par (5.8)
6. calcul de R_{k+1}
7. restauration des contraintes:
   r_{k+1} par (5.4)
8. adaptation de p:
   p_k vérifiant (5.23)-(5.25)

9. recherche sur θ_{p_k} en y_k suivant l'arc (5.2):
   ρ_k par la règle (5.11)-(5.13)
   y_{k+1} par (5.5)
10. linéarisation des contraintes en y_{k+1}:
   $\lambda(y_{k+1})$ par (5.14)
11. mise à jour de H_k:
   H_{k+1} définie positive
12. $k := k+1$
   aller en 5.

A l'étape 11, la définie positivité de H_{k+1} demande l'utilisation d'un schéma de mise à jour concervant cette propriété. Ce schéma peut aussi mettre à jour la matrice G_k et obtenir H_{k+1} par inversion de G_{k+1} supposée non singulière.

V.3 - Convergence globale.

V.3.1 - Théorème.

SI . f et c vérifient (H_0) et (H_3)
. $(x_k) \subset \Omega$, $(y_k) \subset \Omega$ et (H_k) générées par l'algorithme (5.29) avec $R_k = A(x_k)^-$ ou $R_k = A(y_{k-1})^-$ et $0 < \alpha < 1$
. (H_k) bornée et uniformément définie positive

ALORS (i) si (p_k) n'est pas bornée, la suite $(y_k : p_k \neq p_{k-1})$ n'a pas de valeur d'adhérence.

(ii) si (p_k) est bornée, toute valeur d'adhérence de la suite (y_k) est valeur d'adhérence de la suite (x_k) et vérifie les conditions d'optimalité du premier ordre de (1.1)

Preuve.

1) Supposons que (p_k) ne soit pas bornée. Alors, l'ensemble $\{k : k \geq K', p_k \neq p_{k-1}\}$ est non borné. C'est une suite, notons la $(k_i)_{i\geq 0}$. Grâce à (5.23), (5.24), on a pour $i \geq 1$:

$$p_{k_{i-1}} \leq p_{k_i-1} < \|\lambda(y_{k_i})\|_\infty + \bar{p}$$

Comme

$$p_{k_i} \to \infty \text{ quand } i \to \infty,$$

on a

$$||\lambda(y_{k_i})||_\infty \to \infty \text{ quand } i \to \infty$$

λ étant une fonction continue de y, cela montre que la suite considérée en (i) ne peut avoir de valeur d'adhérence.

2) Supposons que (p_k) soit bornée. Alors, d'après (5.25), il existe un indice K_1 et un réel positif p tels que

$$p_k = p \text{ pour tout } k \geq K_1$$

Soit $\bar{y}$ une valeur d'adhérence de (y_k). Comme (ρ_k) et (H_k) sont bornées, il existe des sous-suites (indicées par k_i) telles que lorsque i tend vers l'infini:

$$y_{k_i} \to \bar{y}$$
$$\rho_{k_i} \to \bar{\rho}$$
$$H_{k_i} \to \bar{H}$$

On en déduit successivement les convergences suivantes:

$$g(y_{k_i}) \to \bar{g} := g(\bar{y})$$
$$\lambda(y_{k_i}) \to \lambda(\bar{y})$$
$$q_{k_i} \to \bar{q} := - Z(\bar{y})^- \bar{H} \bar{g}$$
$$x_{k_i+1} \to \bar{y} + \bar{q} \tag{5.30}$$
$$R_{k_i+1} \to \bar{R} \tag{5.31}$$
$$r_{k_i+1} \to \bar{r} := - \bar{R}\, c(\bar{y}+\bar{q})$$
$$y_{k_i+1} \to \bar{y} + \bar{\rho}\, \bar{q} + \bar{\rho}^2\, \bar{r}$$

Dans (5.31), $\bar{R} = A(\bar{y})^-$ ou $A(\bar{y}+\bar{q})^-$ selon le choix de R_k. Pour $k_i \geq K_1$, le paramètre de pénalisation est constant et donc, par construction, la suite (y_k) fait décroître Θ_p (théorème V.2.3). On a

$$\begin{aligned}\Theta_p(y_{k_{i+1}}) &\le \Theta_p(y_{k_i+1})\\ &< \Theta_p(y_{k_i}) - \rho_{k_i}\ \alpha\ (H_{k_i}\ g(y_{k_i})\ ,\ g(y_{k_i})\ )\\ &\quad - \rho_{k_i}^2\ \alpha\ (p - ||\lambda(y_{k_i})||_\infty)\ ||c(y_{k_i})||_1\end{aligned} \tag{5.32}$$

En passant à la limite dans (5.32) lorsque i tend vers l'infini et en tenant compte de la définie positivité des H_k et de (5.23), on trouve:

$$\bar\rho\ (\bar H\ \bar g\ ,\ \bar g) = 0 \tag{5.33}$$

$$\bar\rho\ c(\bar y) = 0 \tag{5.34}$$

Il s'agit de montrer que $\bar g = 0$ et $c(\bar y) = 0$. On désignera par C une constante "générique" indépendante de i et on notera k au lieu de k_i. Soit $0 < \rho < 1$. On cherche à majorer $\Theta_p(y_k + \rho\ q_k + \rho^2\ r_{k+1})$ par des termes semblables à ceux du membre de droite de (5.13). Puisque $c'(y_k).q_k = 0$, on a:

$$c(x_{k+1}) = c(y_k) + O(||q_k||^2) \tag{5.35}$$

Et quelque soit le choix de R_k, on obtient en utilisant (5.35):

$$\begin{aligned}c'(y_k).r_{k+1} &= -\ c(x_{k+1}) + O(||q_k||\ ||r_{k+1}||)\\ &= -\ c(y_k) + O(||q_k||^2) + O(||q_k||\ ||c(y_k)||)\end{aligned} \tag{5.36}$$

$$\begin{aligned}f'(y_k).r_{k+1} &= -\ f'(y_k).A(y_k)^{-}c(x_{k+1}) + O(||q_k||\ ||c(x_{k+1})||)\\ &= (\lambda(y_k),c(y_k)) + O(||q_k||^2) + O(||q_k||\ ||c(y_k)||)\end{aligned} \tag{5.37}$$

Grâce à (5.37), on trouve:

$$\begin{aligned}&f(y_k + \rho\ q_k + \rho^2\ r_{k+1})\\ &\le f(y_k) - \rho\ (H_k g(y_k),g(y_k)) + \rho^2\ ||\lambda(y_k)||_\infty\ ||c(y_k)||_1\\ &\quad + C\ \rho^2\ ||q_k||^2 + C\ \rho^2\ ||q_k||\ ||c(y_k)||_1 + C\ \rho^3\end{aligned} \tag{5.38}$$

Pour le second terme de Θ_p, on calcule, en utilisant (5.36):

$$\begin{aligned}&c^j(y_k + \rho\ q_k + \rho^2\ r_{k+1})\\ &\le (1-\rho^2)\ c^j(y_k) + C\ \rho^2\ ||q_k||^2 + C\ \rho^2\ ||q_k||\ ||c(y_k)||_1 + C\ \rho^3\end{aligned} \tag{5.39}$$

où c^j désigne la j-ième composante de c. En utilisant (5.38), (5.39), l'uniforme définie positivité de (H_k) et $||q_k|| \le C\ ||g(y_k)||$, on trouve finalement, quelque soit ρ dans $]0,1[$:

$$\begin{aligned}&\theta_p(y_k + \rho\, q_k + \rho^2\, r_{k+1})\\ &\le \theta_p(y_k) - \rho\,(1-C_1\rho)\,(H_k g(y_k), g(y_k)) - \rho^2\,(p-||\lambda(y_k)||_\infty)\ ||c(y_k)||_1\\ &\quad + C_2\,\rho^2\ ||q_k||\ ||c(y_k)||_1 + C_3\,\rho^3\end{aligned} \tag{5.40}$$

On exploite (5.33), (5.34) et (5.40) pour montrer que $\bar{g}$ et $c(\bar{y})$ sont nuls et donc que $\bar{y}$ est point stationnaire de (1.1). Montrons d'abord que $\bar{g}$ est nul. Si ce n'est pas le cas, (5.33) et la définie positivité de $\bar{H}$ montrent que $\bar{\rho}$ est nul. D'autre part, comme (H_k) est uniformément définie positive, il existe une constante positive δ et un indice K_2 tels que pour tout indice k supérieur à K_2, on ait:

$$(H_k\ g(y_k)\ ,\ g(y_k)\ ) \ge C\ ||g(y_k)||^2 \ge \delta \tag{5.41}$$

alors, grâce à (5.40), on obtient:

$$\begin{aligned}&\theta_p(y_k + \rho\, q_k + \rho^2\, r_{k+1})\\ &\le \theta_p(y_k) - \rho\,(1 - C_1\rho - C_4\rho/\delta)\,(H_k g(y_k), g(y_k))\\ &\qquad - \rho^2\,\alpha\,(p - ||\lambda(y_k)||_\infty)\ ||c(y_k)||_1\end{aligned} \tag{5.42}$$

où on a tenu compte de (5.41), de (5.23), de $\alpha < 1$ et de ce qu'il existe une constante C_4 indépendante de k telle que

$$C_2\,\rho^2\ ||q_k||\ ||c(y_k)||_1 + C_3\,\rho^3 \le C_4\,\rho^2$$

Soit L un entier naturel tel que

$$\beta^L \le \frac{1-\alpha}{C_1 + C_4/\delta}$$

Alors on a

$$1 - C_1\beta^L - C_4\beta^L/\delta \ge \alpha \tag{5.43}$$

et en prenant $\rho = \beta^L$ dans (5.42), (5.43) montre que (5.13) est vérifiée avec $\rho = \beta^L$, dès que k est supérieur à K_2. Dès lors, la règle de calcul du pas (5.11)-(5.13) donnera pour tout indice k supérieur à K_2:

$$\rho_k \geq \beta^L$$

Ceci contredit le fait que ρ_k converge vers 0. Donc $\bar{g}$ est nul. Supposons à présent que $c(\bar{y})$ ne soit pas nul. Alors d'après (5.34), $\bar{\rho} = 0$. De plus, il existe une constante positive δ et un indice K_3 tels que pour tout indice k supérieur à K_3, on ait:

$$||c(y_k)||_1 \geq \delta \tag{5.44}$$

$$||q_k|| \leq \frac{\bar{p}\,(1-\alpha)}{2C_2} \tag{5.45}$$

(5.45) vient de la convergence de q_k vers 0 que nous venons de démontrer. Utilisons (5.44), (5.45) et (5.23) pour récrire (5.40):

$$\begin{aligned}
&\Theta_p(y_k + \rho\, q_k + \rho^2\, r_{k+1}) \\
&\leq \Theta_p(y_k) - \rho\,(1 - C_1\rho)\,(H_k g(y_k), g(y_k)) \\
&\quad - \rho^2\,(p - ||\lambda(y_k)||_\infty)\,(1 - C_3\rho/(\bar{p}\delta) - (1-\alpha)/2)\,||c(y_k)||_1
\end{aligned} \tag{5.46}$$

Soit L un entier naturel tel que

$$\beta^L \leq \min\left(\frac{1-\alpha}{C_1}, \frac{(1-\alpha)\,\bar{p}\,\delta}{2\,C_3}\right)$$

Alors, on a

$$1 - C_1\beta^L \geq \alpha \tag{5.47}$$

$$1 - C_3\beta^L/(\bar{p}\delta) - (1-\alpha)/2 \geq \alpha \tag{5.48}$$

En prenant $\rho = \beta^L$ dans (5.46), (5.47) et (5.48) montrent que (5.13) est vérifiée avec $\rho = \beta^L$, dès que k est supérieur à K_3. Dès lors, la règle de calcul du pas (5.11)-(5.13) donnera pour tout indice k supérieur à K_3:

$$\rho_k \geq \beta^L$$

Ce qui contredit la convergence de ρ_k vers 0. Donc $c(\bar{y}) = 0$ et $\bar{y}$ est point stationnaire de (1.1).

Pour clore la démonstration, il reste à montrer que $\bar{y}$ est valeur d'adhérence de la suite (x_k). Cela se voit sur (5.30) puisque $\bar{q} = 0$. ●●●

V.4 - Admissibilité asymptotique du pas unité.

Soient $(x_k) \subset \Omega$, $(y_k) \subset \Omega$ et (H_k) générées par l'algorithme modèle global (5.29). $G_k := H_k^{-1}$. Soit (k_i) une sous-suite d'indices. On considérera dans le théorème qui suit des sous-suites avec l'une des quatre propriétés suivantes:

(P_1): $q_{k_i} = O(||r_{k_i+1}||)$ quand $i \to \infty$ (5.49)

(P_2): $(G_{k_i} - G^*)\, Zq_{k_i} = o(||q_{k_i}||)$ quand $i \to \infty$ (5.50)

(P_3): il existe une constante t indépendante de i telle que

$||G_{k_i} - G^*|| \le t$ pour tout indice i. (5.51)

(P_4): $\rho_{k_i-1} = 1$, $0 < \rho_{k_i} < 1$ pour tout indice i et

$q_{k_i} = o(||d_{k_i}||)$ quand $i \to \infty$ (5.52)

(P_2) est une propriété d'approximation de $G^* := Z^{-T}L^*Z^-$ par G_k et rappelle la condition (4.67) du théorème de convergence Q-superlinéaire IV.5.1. En fait, lorsque $\rho_{k-1} = 1$, (P_2) est une propriété plus forte que (4.67) et sera vérifiée pour des sous-suites d'indices dans les algorithmes de la section VI. Le théorème suivant montre que pour les sous-suites (k_i) vérifiant l'une des propriétés (P_1), (P_2) ou (P_3), la règle de détermination de ρ_k (5.11)-(5.13) donnera un pas unité après un nombre fini d'itérations. Lorsque la propriété (P_4) est vérifiée, on ne peut pas conclure dans le même sens dans le cadre de l'algorithme modèle global (5.29). Mais pour les algorithmes concrets de la section VI, le résultat obtenu ci-dessous avec (P_4) permettra de montrer que l'on a également un pas unité après un nombre fini d'itérations. Avoir un pas $\rho_k = 1$ après un nombre fini d'itérations est une propriété importante puisque cela ramène asymptotiquement l'algorithme global (5.29) à l'algorithme local de la section IV. Dans ces conditions une bonne approximation de G^* par G_k conduira à la convergence Q-superlinéaire de la suite (x_k).

V.4.1 - Théorème.

SI . (H_0'), (H_1), (H_2) et (H_3)

. $(x_k) \subset \Omega$, $(y_k) \subset \Omega$ et (H_k) générées par l'algorithme modèle global (5.29) avec $R_k = A(x_k)^-$ ou $R_k = A(y_{k-1})^-$ et $0 < \alpha < 1/2$

. (H_k) bornée, uniformément définie positive

. $y_k \longrightarrow x^*$

. (k_i) une sous-suite d'indices

ALORS (i) si (P_1) ou (P_2) est vérifiée, il existe un indice i_0 tel que $\rho_{k_i} = 1$ pour tout indice i supérieur à i_0

(ii) il existe une constante positive δ telle que si (P_3) est vérifiée avec $t \leq \delta$ alors il existe un indice i_0 tel que $\rho_{k_i} = 1$ pour tout indice i supérieur à i_0

(iii) si (P_4) est vérifiée, alors $r_{k_i+1} = o(||h_{k_i}||\ ||d_{k_i}||)$ quand $i \longrightarrow \infty$

Preuve.

L'uniforme non singularité de (H_k) permet d'utiliser l'équivalence $||h_k|| \sim ||y_k-x^*||$ obtenue en IV.3.5. R_k converge vers A^- qui est injectif. Dès lors, d'après le lemme V.2.5:

$$||r_{k+1}|| \sim ||c(x_{k+1})|| \tag{5.53}$$

Quelque soit le choix de R_k, on obtient alors (c.f.r. (5.37))

$$f'(y_k).r_{k+1} = (\lambda(y_k),c(x_{k+1})) + O(||q_k||\ ||r_{k+1}||)$$

Et en développant $c(x_{k+1})$, on a

$$\begin{aligned} f'(y_k).h_k = &- (H_k g(y_k),g(y_k)) + (\lambda(y_k),c(y_k)) \\ &+ \frac{1}{2}(\lambda(y_k),c''(y_k).q_k^2) + o(||q_k||^2) + O(||q_k||\ ||r_{k+1}||) \end{aligned}$$

On a aussi:

$$f''(y_k).h_k^2 = f''(y_k).q_k^2 + O(||h_k||\ ||r_{k+1}||)$$

D'où finalement

$$\begin{aligned} f(y_k+h_k) &= f(y_k) - (H_k g(y_k),g(y_k)) + (\lambda(y_k),c(y_k)) \\ &\quad + \tfrac{1}{2}\,(L^* q_k,q_k) + o(||q_k||^2) + O(||h_k||\ ||r_{k+1}||) \end{aligned} \tag{5.54}$$

D'autre part, en remarquant que $c'(y_k).q_k = 0$ et en notant $s_k^t := y_k + tq_k$, on obtient comme en (5.36) et quelque soit le choix de R_k:

$$c'(y_k).h_k = -\,c(y_k) - \tfrac{1}{2}\int_0^1 c''(s_k^t).q_k^2\,dt + O(||q_k||\ ||r_{k+1}||)$$

Ce qui permet d'obtenir, en notant $z_k^t := y_k + th_k$:

$$\begin{aligned} c(y_k+h_k) &= \int_0^1 \tfrac{1}{2}\,[c''(z_k^t)-c''(s_k^t)].q_k^2 + c''(z_k^t).(q_k)(r_{k+1}) \\ &\quad + \tfrac{1}{2}\,c''(z_k^t).r_{k+1}^2\ \ dt \end{aligned}$$

$$c(y_k+h_k) = o(||q_k||^2) + O(||h_k||\ ||r_{k+1}||) \tag{5.55}$$

D'après le théorème V.3.1, l'existence d'une valeur d'adhérence pour (y_k) et (5.24) impliquent que p_k est modifié un nombre fini de fois. On supposera que $p_k = p$, pour tout indice k. On notera

$$Y_k := -\,(H_k g(y_k),g(y_k)) - (p-||\lambda(y_k)||_\infty)\ ||c(y_k)||_1$$

$$X_k := \Theta_p(y_k+h_k) - \Theta_p(y_k) - \alpha\, Y_k$$

$Y_k \le 0$. Il s'agit de voir dans quel cas X_k est négatif. Grâce à (5.54) et (5.55), on obtient:

$$\begin{aligned} X_k &\le (\tfrac{1}{2}-\alpha)\ Y_k - \tfrac{1}{2}\,(H_k g(y_k),g(y_k)) + \tfrac{1}{2}\,(L^* q_k,q_k) \\ &\quad + o(||q_k||^2) + O(||h_k||\ ||r_{k+1}||) \end{aligned}$$

Comme $g(y_k) = G_k\, Z\, q_k + o(||q_k||)$ et $q_k = Z^-\, Z\, q_k + o(||q_k||)$, on obtient

$$\begin{aligned} X_k &\le (\tfrac{1}{2}-\alpha)\ Y_k - \tfrac{1}{2}\,((G_k-G^*)Zq_k,Zq_k) \\ &\quad + o(||q_k||^2) + O(||h_k||\ ||r_{k+1}||) \end{aligned} \tag{5.56}$$

On exploite cette inégalité pour obtenir les résultats recherchés. Supposons qu'il existe une sous-suite L d'indices l pour lesquels $\rho_l < 1$, c'est-à-dire que l'on a $X_l \geq 0$. En tenant compte de ce que $\alpha < 1/2$, de l'uniforme définie positivité de (H_k) et de (5.23), (5.56) donne:

$$\|q_l\|^2 \leq - C_1 \, ((G_l-G^*)Zq_l, Zq_l) + o(\|q_l\|^2) + O(\|h_l\| \; \|r_{l+1}\|) \tag{5.57}$$

$$\|c(y_l)\| \leq - C \, ((G_l-G^*)Zq_l, Zq_l) + o(\|q_l\|^2) + O(\|h_l\| \; \|r_{l+1}\|) \tag{5.58}$$

1) Pour prouver (i), il suffit de montrer que l'on ne peut avoir ni (P_1) ni (P_2) pour la sous-suite L. Supposons que (P_1) soit vérifiée avec $k_i=l$, $l \in L$. Grâce à (5.58), on obtient:

$$\begin{aligned} &\|q_l\| \leq C \, \|r_{l+1}\| \\ &\|c(y_l)\| \leq C \, \|r_{l+1}\|^2 \end{aligned} \tag{5.59}$$

Or, d'après (5.35) et l'équivalence (5.53), on a

$$r_{k+1} = O(\|c(y_k)\| + \|q_k\|^2) \tag{5.60}$$

On obtient alors, quelque soit $l \in L$:

$$\|r_{l+1}\| \leq C \, \|r_{l+1}\|^2$$

ce qui est impossible, puisque lorsque $l \longrightarrow \infty$, $r_{l+1} \longrightarrow 0$.

Supposons à présent que (P_2) soit vérifiée avec $k_i=l$, $l \in L$. (5.57) donne:

$$\|q_l\|^2 \leq \eta_l \, \|q_l\|^2 + C \, \|q_l\| \; \|r_{l+1}\| + C \, \|r_{l+1}\|^2$$

où η_l converge vers 0 lorsque l tend vers l'infini. En utilisant l'inégalité de Young ($ab \leq a^2/2 + b^2/2$) sur le produit $(\|q_l\|)(C\|r_{l+1}\|)$, on voit que $q_l = O(\|r_{l+1}\|)$ pour l assez grand. Comme précédemment, cela conduit à une contradiction.

2) Soit $\delta < (C_1\|Z\|^2)^{-1}$, C_1, donnée en (5.57), étant une constante indépendante de k donc de $l \in L$. Pour prouver (ii), il suffit de montrer que la sous-suite L ne peut pas vérifier (P_3) avec $t \leq \delta$. Supposons que (P_3) soit vérifiée avec $t \leq \delta$ et $k_i=l$, $l \in L$. Alors (5.57) donne:

$$\|q_l\|^2 \leq t \; C_1 \; \|Z\|^2 \; \|q_l\|^2 + \eta_l \; \|q_l\|^2 + C \; \|h_l\| \; \|r_{l+1}\|$$

Comme $t \leq \delta$, on déduit de cette inégalité, l'existance de constantes C_2 et C_3 indépendantes de l telles que pour l assez grand on ait:

$$||q_l||^2 \leq C_2 \; ||q_l|| \; ||r_{l+1}|| + C_3 \; ||r_{l+1}||^2$$

L'utilisation de l'inégalité de Young pour le premier terme du second membre montre que l'on a encore $q_l = O(||r_{l+1}||)$ pour l assez grand. Ceci, comme dans le premier cas, conduit à une contradiction.

3) Afin de prouver (iii), supposons que (P_4) soit vérifiée avec $k_i=1$, $l \in L$. Comme $\rho_{l-1} = 1$, on a, d'après les estimations (4.36) et (4.31) et l'équivalence (4.40):

$$r_{l+1} = o(||d_l||) \tag{5.61}$$

De (P_4),(5.57),(5.58) et (5.61), on obtient:

$$||q_l||^2 \leq \eta_l \; ||h_l|| \; ||d_l||$$
$$||c(y_l)|| \leq \eta_l \; ||h_l|| \; ||d_l||$$

où η_l converge vers 0 lorsque l tend vers l'infini. Ces inégalités et (5.60) conduisent au résultat (iii). ●●●

VI - MISE A JOUR DE LA MATRICE REDUITE

VI.1 - Introduction et hypothèse supplémentaire.

On va étudier dans cette section la convergence d'algorithmes concrets calqués sur l'algorithme modèle global (5.29) en ce sens que l'on se fixe les matrices R_k et H_k afin que sous des conditions raisonnables, on obtienne la convergence Q-superlinéaire.

En ce qui concerne l'opérateur R_k, on prendra soit $R_k = A(x_k)^-$, soit $R_k = A(y_{k-1})^-$. Dès lors, si on suppose la convergence de y_k vers x^*, point stationnaire de (1.1), on a

$$R_k \longrightarrow A^- \text{ quand } k \longrightarrow \infty \tag{6.1}$$

Ceci règle le problème de l'approximation de A^-, c'est-à-dire que les clauses (4.64) et (4.66) du théorème de convergence Q-superlinéaire IV.5.1 sont vérifiées, ainsi que les conditions sur R_k utilisées dans les théorèmes de la section V.

D'autre part, on cherche à ce que G_k soit proche de G^* dans le sens des conditions (4.65) et (4.67) du théorème IV.5.1, ou mieux, dans le sens de la condition (5.50) du théorème V.4.1. Selon ces théorèmes, avoir G_k proche de G^* est intéressant pour obtenir la convergence Q-superlinéaire de la suite (x_k) (théorème IV.5.1). Mais pour cela, il faut que ρ_k soit égal à 1. C'est encore G_k proche de G^* qui permettra d'avoir un pas unité après un nombre fini d'itérations (théorème V.4.1). On voit donc qu'il est nécessaire de mettre à jour la matrice G_k même lorsque ρ_k diffère de 1 et par conséquent d'étudier cette mise à jour dans le cadre de la méthode globale (5.29). C'est ce que nous allons faire. Les notations de cette section seront celles de la section V (c.f.r. (5.3)-(5.10)).

Pour la mise à jour de G_k, on utilisera une méthode de quasi-Newton. C'est-à-dire que l'on va construire G_{k+1} à partir de G_k en imposant à G_{k+1} de vérifier une équation dite de quasi-Newton ou de la sécante:

$$\gamma_k = G_{k+1}\,\sigma_k \qquad (6.2)$$

où, γ_k et σ_k sont des vecteurs de $\mathbb{R}^{n-m}$. On impose à G_{k+1} de vérifier (6.2) parce que l'on va s'arranger pour que γ_k soit assez proche de $G^*\sigma_k$ (voir plus loin). Dès lors, avec (6.2), G_{k+1} aura sur σ_k un effet assez semblable à celui de G^*. Toutefois, les (n-m) équations de (6.2) ne suffisent pas pour déterminer G_{k+1} qui est d'ordre (n-m), même si elle est symétrique. On doit donc s'imposer des conditions supplémentaires. Celles-ci conduisent en général à faire intevenir G_k dans le calcul de G_{k+1}. Si on suppose que la formule de mise à jour de G_k porte sur son inverse H_k, on peut de façon assez générale formaliser ce calcul par:

$$H_{k+1} := U\,(H_k,\ \gamma_k,\ \sigma_k) \qquad (6.3)$$

Enfin, il est apparu clairement dans la section V que le caractère défini positif de H_k (ou de G_k) est essentiel à l'obtention de la convergence globale et à l'admissibilité asymptotique du pas unité. On supposera donc que H_0 est symétrique, définie positive et on prendra dans (6.3) une formule de mise à jour de H_k qui transmet ces deux propriétés à H_{k+1}.

Dans cette section, on fera systématiquement l'hypothèse suivante:

(H_2') | G^* est définie positive

qui est plus forte que (H_2). L'ensemble des hypothèses (H_1) et (H_2') constitue des conditions suffisantes pour que x^* soit un minimum local strict pour le problème (1.1).

VI.2 - Formule de mise à jour.

VI.2.1 - Formule BFGS.

Les schémas de mise à jour de matrices sous la forme (6.3) et sur base de l'équation de la sécante (6.2) ont été introduits pour les méthodes à métrique variable en optimisation sans contrainte (voir la revue de Dennis, Moré (1977)). Les théorèmes de convergence locale et Q-superlinéaire que nous établirons le seront pour une formule de mise à jour (6.3) particulière à savoir la formule BFGS (Broyden (1969), Fletcher (1970), Goldfarb (1970), Shanno (1970)) qui construit H_{k+1} par addition à H_k de deux matrices de rang 1:

$$H_{k+1} := H_k + \frac{(\sigma_k - H_k\gamma_k)\sigma_k^T + \sigma_k(\sigma_k - H_k\gamma_k)^T}{(\gamma_k,\sigma_k)} - \frac{(\gamma_k,\sigma_k - H_k\gamma_k)}{(\gamma_k,\sigma_k)^2}\,\sigma_k\sigma_k^T . \tag{6.4}$$

Succinctement, on représentera cette formule par

$$H_{k+1} := \mathrm{BFGS}\ (H_k,\ \gamma_k,\ \sigma_k) \tag{6.5}$$

La formule BFGS est bien définie dès que (γ_k,σ_k) est non nul. H_{k+1} sera symétrique si H_k l'est. Si H_k est symétrique définie positive, H_{k+1} sera définie positive si et seulement si (γ_k,σ_k) est positif (Dennis, Moré (1977)). Dès lors on gardera la définie positivité de H_k si la mise à jour ne se fait que si (γ_k,σ_k) est positif.

On désignera par $||.||_F$, la norme matricielle de Frobenius définie sur $L(\mathbb{R}^{n-m})$ par:

$$Q \in L(\mathbb{R}^{n-m}) \quad \rightarrow \quad ||Q||_F := (\ \mathrm{tr}(Q^T\ Q)\ )^{1/2}$$

Si M est une matrice d'ordre n-m non singulière, on définit sur $L(\mathbb{R}^{n-m})$ la norme matricielle $|||.|||$ par

$$Q \in L(\mathbb{R}^{n-m}) \quad \rightarrow \quad |||Q||| := ||MQM||_F \tag{6.6}$$

VI.2.2 - <u>Lemme.</u>

SI . M matrice d'ordre n-m, symétrique, non singulière

. γ et σ sont des vecteurs de $\mathbb{R}^{n-m}$ tels que $(\gamma,\sigma) \neq 0$ et $||M\,\sigma - M^{-1}\,\gamma|| \le \frac{1}{3}\,||M^{-1}\,\gamma||$ (6.7)

. H et A deux matrices d'ordre n-m, symétriques

. $\bar{H} := \mathrm{BFGS}\ (H,\ \gamma,\ \sigma)$

ALORS
$$|||\bar{H} - A||| \le \left(1 - \theta^2 + \frac{15}{4}\,\frac{||M\,\sigma - M^{-1}\,\gamma||}{||M^{-1}\,\gamma||}\right)|||H - A||| + 2\,(1 + 2(n-m)^{1/2})\,||M||_F\,\frac{||\sigma - A\,\gamma||}{||M^{-1}\,\gamma||} \tag{6.8}$$

où

$$\theta := \begin{cases} \dfrac{(3)^{1/2}}{4} \dfrac{||M\ (H-A)\ \gamma||}{|||H-A|||\ ||M^{-1}\ \gamma||} & \text{si } H \neq A \\ 0 \quad \text{sinon} & \end{cases} \tag{6.9}$$

Ce lemme se déduit aisément du lemme 5.2 de Broyden, Dennis, Moré (1973) en prenant dans celui-ci $\beta = 1/3$ et en utilisant l'inégalité $(1-2\theta^2)^{1/2} \leq 1-\theta^2$. Par la suite, on prendra $M^{-2} := A := H^*$. On voit alors dans (6.8) que l'approximation de H^* par $\bar{H}$ est contrôlée par l'écart entre H^* et H et par la manière dont γ et σ sont choisis par rapport à H^*.

Le lemme suivant est un résultat de Powell (1971) (théorème 3) donnant des conditions sur γ_k et σ_k pour que les suites (H_k) et (H_k^{-1}) soient bornées lorsque (H_k) est générée par la formule BFGS (6.4).

VI.2.3 - Lemme.

SI
- A une matrice d'ordre n-m, symétrique, définie positive
- (γ_k) et (σ_k) deux suites de $\mathbb{R}^{n-m}$
- $(\gamma_k,\sigma_k) > 0$ pour tout indice k (6.10)
- $\left(\dfrac{||\sigma_k - A\ \gamma_k||}{||\gamma_k||}\right)$ sommable (6.11)
- H_0 une matrice d'ordre n-m, symétrique, définie positive

ALORS la formule BFGS génère à partir de H_0, une suite de matrices symétriques (H_k), bornée et uniformément définie positive

Preuve.

On peut récrire la formule (6.4) comme suit:

$$H_{k+1} = \left[I - \frac{\sigma_k\ \gamma_k^T}{(\gamma_k,\sigma_k)} \right] H_k \left[I - \frac{\gamma_k\ \sigma_k^T}{(\gamma_k,\sigma_k)} \right] + \frac{\sigma_k\ \sigma_k^T}{(\gamma_k,\sigma_k)} \tag{6.12}$$

Comme (γ_k,σ_k) est positif, H_{k+1} est non singulière et son inverse G_{k+1} est donné par la formule

$$G_{k+1} = G_k - \frac{G_k \sigma_k \sigma_k^T G_k}{(\sigma_k, G_k \sigma_k)} + \frac{\gamma_k \gamma_k^T}{(\gamma_k, \sigma_k)}$$

Il s'agit de montrer que (H_k) et (G_k) sont bornées. Soit $A^{1/2}$ la racine carrée définie positive de A et

$$Q_k := I - \frac{A^{1/2} \gamma_k \gamma_k^T A^{1/2}}{(A\gamma_k, \gamma_k)} + \frac{A^{-1/2} \sigma_k \sigma_k^T A^{-1/2}}{(\gamma_k, \sigma_k)}$$

Comme A est définie positive et $(\gamma_k, \sigma_k) > 0$, Q_k est définie positive. On définit alors z_k, J_k et R_k par:

$$z_k := Q_k^{1/2} A^{1/2} \gamma_k = Q_k^{-1/2} A^{-1/2} \sigma_k$$

$$J_k := Q_k^{1/2} A^{1/2} G_k A^{1/2} Q_k^{1/2}$$

$$R_k := J_k - \frac{J_k z_k z_k^T J_k}{(J_k z_k, z_k)}$$

Ce qui permet d'écrire

$$Q_k^{1/2} A^{1/2} G_{k+1} A^{1/2} Q_k^{1/2} = R_k + \frac{z_k z_k^T}{||z_k||^2}$$

Comme $R_k z_k = 0$ et $||R_k||_2 \leq ||J_k||_2$, cette égalité donne

$$||Q_k^{1/2} A^{1/2} G_{k+1} A^{1/2} Q_k^{1/2}||_2 \leq \max(1, ||J_k||_2)$$

ou encore, en notant $m_k := \max(1, ||Q_k^{-1}||_2) \max(1, ||Q_k||_2)$

$$||A^{1/2} G_{k+1} A^{1/2}||_2 \leq m_k \max(1, ||A^{1/2} H_k A^{1/2}||_2)$$

De manière analogue, on obtient à partir de (6.12)

$$||A^{-1/2} H_{k+1} A^{-1/2}||_2 \leq m_k \max(1, ||A^{-1/2} H_k A^{-1/2}||_2)$$

Dès lors, il suffit de montrer que $\Pi\, m_k < +\infty$. Pour cela, on cherche une majoration de $||Q_k||_2$ et $||Q_k^{-1}||_2$. On écrit

$$Q_k - I = \frac{A^{-1/2} (\sigma_k - A\gamma_k) \sigma_k^T A^{-1/2}}{(\gamma_k, \sigma_k)} + \frac{A^{1/2} \gamma_k (\sigma_k - A\gamma_k)^T A^{-1/2}}{(\gamma_k, \sigma_k)}$$

$$+ \frac{(A\gamma_k,\gamma_k) - (\gamma_k,\sigma_k)}{(\gamma_k,\sigma_k)\ (A\gamma_k,\gamma_k)}\ A^{1/2}\ \gamma_k\ \gamma_k^T\ A^{1/2}$$

Soit C_0^2 la plus petite valeur propre de A. D'après (6.10) et (6.11), il existe des constantes positives C_1 et C_2 telles que $(\gamma_k,\sigma_k) \geq C_1\ ||\gamma_k||^2$ et $||\sigma_k|| \leq C_2\ ||\gamma_k||$. En notant $\varepsilon_k := ||\sigma_k - A\gamma_k||/||\gamma_k||$ et $C_3 := C_2/C_0^2C_1 + ||A||_2^{1/2}/C_0C_1 + 1/C_1$, la dernière équation donne

$$||Q_k - I||_2 \leq C_3\ \varepsilon_k \tag{6.13}$$

Soit $C_4 > C_3$. Pour k assez grand, disons $k \geq K_1$, on a $C_3\varepsilon_k \leq 1 - C_3/C_4$. Alors, on a grâce à (6.13)

$$||Q_k||_2 \leq 1 + C_3\ \varepsilon_k$$

$$||Q_k^{-1}||_2 \leq 1 + C_4\ \varepsilon_k \quad \text{pour } k \geq K_1$$

Ces deux inégalités et (6.11) permettent de montrer que $\Pi\ m_k < +\infty$. ●●●

VI.2.4 - Lemme.

SI
- C_1, C_2 et C_3 trois constantes non négatives
- (δ_k), (θ_k) et (χ_k) trois suites de réels non négatifs telles que pour tout indice k on ait:

$$\delta_{k+1} \leq (1 - C_1\ \theta_k + C_2\ \chi_k)\ \delta_k + C_3\ \chi_k \tag{6.14}$$

- $$\sum_{k=0}^{\infty} \chi_k < +\infty \tag{6.15}$$

ALORS

(i) $\delta_k \longrightarrow \delta$ quand $k \longrightarrow \infty$

(ii) si $C_1 > 0$ alors soit $\delta = 0$ soit $\theta_k \longrightarrow 0$

Ce lemme se déduit aisément du lemme 3.3 de Dennis, Moré (1974).

VI.3 - Deux algorithmes.

VI.3.1 - Choix des vecteurs γ_k et σ_k.

Les formules (6.3) et (6.4) montrent que, étant donné H_k, la "qualité" de H_{k+1} dépendra du choix des vecteurs γ_k et σ_k. On l'a vu, un bon choix consisterait à prendre des couples (γ_k, σ_k) vérifiant la relation:

$$\sigma_k = H^* \gamma_k \tag{6.16}$$

et à engendrer $\mathbb{R}^{n-m}$ par les vecteurs σ_k. En pratique, (6.16) n'est pas vérifiée, mais l'étude des méthodes de quasi-Newton montre qu'il est essentiel qu'elle le soit au premier ordre, c'est-à-dire que l'on ait:

$$\sigma_k = H^* \gamma_k + o(||\gamma_k||) \tag{6.17}$$

La condition (6.17) apparaît en effet en (6.7) (avec $M = (H^*)^{-1/2}$), en (6.8) et en (6.11) (avec $A = H^*$). H^* étant non singulière, cette relation montre que l'on a $||\gamma_k|| \sim ||\sigma_k||$. Dès lors, (6.17) est équivalente à

$$\gamma_k = G^* \sigma_k + o(||\sigma_k||) \tag{6.18}$$

où G^* est l'inverse de H^*. Comme $G^* = Z^{-T} L^* Z^-$, les relations (6.18) et (2.31) nous renseignent sur la manière de choisir γ_k et σ_k. γ_k doit être la différence de deux gradients réduits g calculés en des points, disons b_k et a_k, tels que $(b_k-a_k)/||b_k-a_k||$ appartienne asymptotiquement à l'image de Z^-. On choisit alors σ_k de la forme $Z(b_k-a_k)$.

VI.3.2 - L'algorithme ALG1.

Supposons que (y_k) converge vers x^*. Un premier choix de γ_k et σ_k est immédiat. Etant donné que $q_k = x_{k+1} - y_k$ est dans l'image de $Z(y_k)^-$ (voir formule (5.3)), on choisit:

$$\gamma_k := g(x_{k+1}) - g(y_k) \tag{6.19}$$

$$\sigma_k := Z(x_{k+1})\, q_k \tag{6.20}$$

Dans (6.19), le calcul du gradient réduit en x_{k+1} nécessite la linéarisation des contraintes en ce point. Alors $Z(x_{k+1})$ est disponible et on le choisit de préférence à $Z(y_k)$ dans (6.20) parce qu'il est sensé être plus proche de $Z := Z(x^*)$. Mais ce n'est pas fondamental. On obtient alors, grâce à l'identité (2.31) et à l'hypothèse de régularité (H_0') (voir paragraphe V.1):

$$\gamma_k = Z^{-T} L^* q_k + o(||q_k||) \tag{6.21}$$

D'autre part, d'après la définition (5.3) de q_k, on voit que l'on a $||q_k|| \sim ||Zq_k||$. Dès lors, σ_k défini en (6.20) vérifie $||\sigma_k|| \sim ||q_k||$, et on a:

$$q_k = Z(y_k)^- Z(y_k) q_k = Z^- \sigma_k + o(||\sigma_k||) \tag{6.22}$$

(6.21) et (6.22) montrent que (6.18) est vérifiée.

Le choix (6.19)-(6.20) de γ_k et σ_k constitue une simplification de ce qu'ont proposé Coleman, Conn (1984). Il présente toutefois un inconvénient. Dans l'algorithme modèle (5.29), le gradient réduit g de f n'est connu qu'en y_k et pas en x_{k+1}. Dès lors, l'utilisation de (6.19) nécessite une relinéarisation des contraintes en x_{k+1}. Cette linéarisation supplémentaire peut apparaître comme le coût de la réduction de n à n-m de la taille de la matrice à mettre à jour.

<u>Algorithme ALG1</u>: (6.23)

1. choix de $x_0 \in \Omega$ et H_0 symétrique définie positive
   $k := 0$
2. linéarisation des contraintes en x_0
3. restauration des contraintes:
   $y_0 := x_0 - A(x_0)^- c(x_0)$ (6.24)

4. linéarisation des contraintes en y_k:
   $\lambda(y_k)$ par (5.14), $g(y_k)$ par (2.29)
5. minimisation:
   q_k par (5.3), x_{k+1} par (5.8)
6. linéarisation des contraintes en x_{k+1}:
   $g(x_{k+1})$ par (2.29)

$r_{k+1} := - A(x_{k+1})^{-} c(x_{k+1})$ (6.25)

7. mise à jour de H_k:

   γ_k par (6.19), σ_k par (6.20)

   <u>si</u> $(\gamma_k, \sigma_k) > 0$ <u>alors</u> $H_{k+1} :=$ BFGS $(H_k, \gamma_k, \sigma_k)$

   <u>sinon</u> $H_{k+1} := H_k$

8. adaptation de p:

   p_k vérifiant (5.23)-(5.25)

9. recherche sur Θ_{p_k} en y_k suivant l'arc (5.2):

   ρ_k par la règle (5.11)-(5.13) avec $0 < \alpha < 1/2$

   y_{k+1} par (5.5)

10. $k := k+1$

    aller en 4.

On voit en (6.25) que l'on a pris $R_{k+1} = A(x_{k+1})^{-}$. Ceci est possible car les contraintes sont linéarisées en x_{k+1} $(A(x_{k+1}) = c'(x_{k+1}))$. Celles-ci le sont donc deux fois (en y_k et x_{k+1}, étapes 4 et 6) par itération (étapes 4 à 10). La matrice H_k n'est remise à jour que si (γ_k, σ_k) est positif, ce qui garantit la définie positivité de H_{k+1}.

VI.3.3 - <u>Autre choix de γ_k et σ_k.</u>

Le fait de devoir relinéariser les contraintes en x_{k+1} peut dans certain cas constituer un surcoût important: d'abord pour le calcul des éléments constituant $A(x_{k+1})$ et ensuite pour celui de l'inverse à droite $A(x_{k+1})^{-}$. Par exemple, dans l'application traitée dans Blum, Gilbert, Thooris (1985) où il s'agit d'identifier des paramètres (les n-m variables de contrôle) dans une équation aux dérivées partielles (les m contraintes), le temps de calcul pour linéariser les contraintes est environ 10 fois plus élevé que le temps de calcul d'une itération interne (qui correspond ici au calcul du pas de minimisation q_k). Il est donc important de voir s'il est possible de se passer de cette linéarisation supplémentaire en n'utilisant que l'information disponible dans l'algorithme modèle global, à savoir les gradients réduits de f en y_k et en y_{k+1}. Cette fois, on choisit:

$$\gamma_k := g(y_{k+1}) - g(y_k) \tag{6.26}$$

$$\sigma_k := Z(y_{k+1})\, v_k \tag{6.27}$$

où, $v_k := y_{k+1} - y_k = \rho_k q_k + \rho_k^2 r_{k+1}$. Si on suppose la convergence de y_k vers x^*, on a, grâce à (2.31):

$$\gamma_k = Z^{-T} L^* v_k + o(||v_k||) \tag{6.28}$$

Mais (6.27) montre qu'en général, il n'existe pas de constante C, indépendante de k, telle que $||v_k|| \leq C\,||\sigma_k||$ ($Z(y_{k+1})$ n'est pas injectif). Dès lors, (6.28) ne conduira en général pas à (6.18). Le lemme suivant permet de dégager des situations dans lesquelles cela se produit.

VI.3.4 - Lemme.

SI . (H_0), (H_1), (H_2) et (H_3)

. $(x_k) \subset \Omega$, $(y_k) \subset \Omega$ et (H_k) générées par l'algorithme modèle global (5.29) avec (6.1)

. (H_k) bornée et uniformément définie positive

. $y_k \to x^*$

ALORS les clauses suivantes sont équivalentes:

(i) $\rho_k^2\, c(y_k) = o(||\rho_k q_k||)$ (6.29)

(ii) $\rho_k^2\, r_{k+1} = o(||\rho_k q_k||)$ (6.30)

(iii) $v_k = Z^- Z\, v_k + o(||\rho_k q_k||)$ (6.31)

DE PLUS ces clauses impliquent:

(iv) $||v_k|| \sim ||Zv_k|| \sim ||\rho_k q_k||$

Preuve.

Comme (H_k) est bornée et uniformément non singulière, on utilisera systématiquement l'équivalence (4.41). Les relations (6.1) et (4.36) d'une part et (4.29) d'autre part donnent:

$$A^- c(y_k) + r_{k+1} = o(||h_k||) \tag{6.32}$$

Si (i) est vérifiée, (6.32) montre que $\rho_k r_{k+1} = o(||q_k||) + \rho_k o(||h_k||)$. Comme $0 < \rho_k \leq 1$, on en déduit (ii). Inversément, si (ii) est vérifiée, (6.32) montre que $\rho_k A^- c(y_k) = o(||q_k||)$. On en déduit (i), grâce à l'injectivité de A^-. En multipliant (6.32) par Z, on trouve:

$$Z\ r_{k+1} = o(||h_k||) \tag{6.33}$$

D'après (4.11), $q_k = Z^- Zq_k + o(||q_k||)$ et donc, avec (6.33), on obtient:

$$Z^-\ Z\ v_k = \rho_k\ q_k + \rho_k\ o(||q_k||) + \rho_k^2\ o(||h_k||) \tag{6.34}$$

$$v_k = Z^-\ Z\ v_k + \rho_k^2\ r_{k+1} + \rho_k\ o(||q_k||) + \rho_k^2\ o(||r_{k+1}||)$$

Cette dernière relation permet de montrer l'équivalence entre (ii) et (iii). Montrons à présent que (i)-(iii) implique (iv). La relation (ii) montre que $v_k = \rho_k q_k + \rho_k o(||q_k||)$ et donc $||v_k|| \sim ||\rho_k q_k||$. Alors (iii) donne $v_k = Z^- Zv_k + o(||v_k||)$ ce qui grâce à l'injectivité de Z^- montre que $||v_k|| \sim ||Zv_k||$. ●●●

VI.3.5 - L'algorithme ALG2.

Supposons que les clauses (i)-(iv) du lemme VI.3.4 soient vérifiées. Alors σ_k donné par (6.27) s'écrit:

$$\sigma_k = Z\ v_k + o(||Zv_k||)$$

D'où l'on déduit

$$||\sigma_k|| \sim ||Zv_k|| \sim ||v_k|| \sim ||\rho_k q_k|| \tag{6.35}$$

$$\sigma_k = Z\ v_k + o(||\sigma_k||) \tag{6.36}$$

Et en utilisant (6.28), (6.31), (6.35) et (6.36), on obtient (6.18). Cette relation (6.18) sera donc vérifiée si on a par exemple:

$$\rho_k\ r_{k+1} = o(||q_k||) \tag{6.37}$$

Dans le cas contraire, r_{k+1} reste grand devant q_k et on peut espérer que le fait de mettre à jour H_k ou pas ne soit pas trop important pour la convergence de l'algorithme. Ces remarques nous conduisent à la définition d'algorithmes où la mise à jour de H_k, ne se faisant pas à chaque itération, est conditionnée à la réalisation d'un critère basé sur (6.37). C'est-à-dire que la matrice réduite H_k sera mise à jour à l'itération k, si on a

$$\rho_k \; ||r_{k+1}|| \leq \mu'_k \; ||q_k|| \tag{6.38}$$

où (μ'_k) est une suite convergeant vers 0. Le problème revient à choisir adéquatement cette suite (μ'_k), si possible en fonction de l'information disponible à l'itération courante de telle sorte que pour les indices k où le critère n'est pas vérifié, on ait quand même la relation de convergence Q-superlinéaire: $x_{k+1}-x^* = o(||x_k-x^*||)$.

Récemment Nocedal, Overton (1985) ont étudié la mise à jour de la matrice réduite de la méthode de Gabay (1982,b). Ils utilisent également un critère de mise à jour semblable à (6.38) mais où la suite (μ'_k) est donnée à priori: $\mu'_k = \eta/(1+k)^{1+\nu}$. Les suites générées convergent alors Q-superlinéairement en 2 pas (on sait, c.f.r. l'exemple de Byrd (1985), qu'en général on ne peut pas obtenir un meilleur résultat pour l'algorithme de Gabay). Toutefois leur choix de μ'_k semble souffrir de deux inconvénients. D'abord, leur étude étant locale ((x_0,H_0) est supposé proche de (x^*,H^*)), la suite générée (x_k) converge Q-linéairement (en 2 pas) et c'est ce fait qui permet de prendre une loi à priori pour (μ'_k). Nos théorèmes de convergence Q-superlinéaire supposent une hypothèse légèrement plus faible que la convergence Q-linéaire. Ensuite, si ν peut être choisi arbitrairement, il n'en va pas de même pour η qui doit être choisi assez petit afin d'assurer la convergence Q-linéaire. Mais si η est pris trop petit, la matrice réduite est rarement mise à jour et la convergence est lente. Ici, la suite (μ'_k) ne sera pas donnée à priori mais calculée en fonction de l'information disponible à l'itération courante et la convergence Q-superlinéaire de (x_k) ne dépendra pas d'un choix de constante.

On désignera par K, l'ensemble des indices k pour lesquels il y a mise à jour de H_k en utilisant (6.4) et par K^c son complémentaire. Par exemple pour l'algorithme ALG1, on aura

$$K = \{ \; k \in N \; : \; (\gamma_k,\sigma_k) > 0 \; \}$$

Pour tout indice k, on définit (k-) par

$$\begin{aligned} &k- := 0 \;\; \underline{\text{si}} \;\; i \in K \Rightarrow i \geq k \\ &k- := \max \{ \; i \in K : i < k \; \} \;\; \underline{\text{sinon}} \end{aligned} \tag{6.39}$$

C'est donc le dernier indice précédent k pour lequel il y a eu mise à jour de H_k par (6.4). On notera également:

$(k--) := (k-)-$

Algorithme ALG2:

1. choix de $x_0 \in \Omega$, H_0 symétrique définie positive et $\mu_0 > 0$
   $k := 0$
2. linéarisation des contraintes en x_0
3. restauration des contraintes:
   $y_0 := x_0 - A(x_0)^- c(x_0)$ (6.40)
4. linéarisation des contraintes en y_0:
   $\lambda(y_0)$ par (5.14), $g(y_0)$ par (2.29)

5. minimisation:
   q_k par (5.3), x_{k+1} par (5.8)
6. restauration des contraintes:
   $r_{k+1} := - A(y_k)^- c(x_{k+1})$ (6.41)
7. adaptation de p:
   p_k vérifiant (5.23)-(5.25)
8. recherche sur Θ_{p_k} en y_k suivant l'arc (5.2):
   ρ_k par la règle (5.11)-(5.13) avec $0 < \alpha < 1/2$
   y_{k+1} par (5.5)
9. linéarisation des contraintes en y_{k+1}:
   $\lambda(y_{k+1})$ par (5.14), $g(y_{k+1})$ par (2.29)
10. mise à jour de H_k:
   γ_k par (6.26), σ_k par (6.27)
   <u>si</u> ($(\gamma_k, \sigma_k) > 0$ <u>et</u>
   $\rho_k \cdot \|r_{k+1}\| \le \mu_k \, \|h_{k--}\| \, \|q_k\|$) (6.42)
   <u>alors</u> $H_{k+1} :=$ BFGS $(H_k, \gamma_k, \sigma_k)$
   <u>sinon</u> $H_{k+1} := H_k$
11. adaptation de μ_k:
   <u>si</u> ($(\gamma_k, \sigma_k) \le 0$ <u>et</u> $\sigma_k \ne 0$ <u>et</u> (6.42))
   <u>alors</u> $\mu_{k+1} := \mu_k / 2$
   <u>sinon</u> $\mu_{k+1} := \mu_k$
12. $k := k+1$

aller en 5.

L'algorithme ALG2 ne nécessite la linéarisation des contraintes qu'une fois (en y_{k+1}, étape 9) par itération (étapes 5 à 12). La restauration des contraintes en x_k se fait avec l'opérateur $R_{k+1} = A(y_k)^-$ (c.f.r. (6.41)), sauf initialement (c.f.r. (6.40)). On voit à l'étape 10 que la matrice H_k n'est pas mise à jour à chaque itération. Il faut pour cela d'une part que (γ_k, σ_k) soit positif, ce qui permet de conserver la définie positivité des H_k, et d'autre part que le critère (6.42) soit vérifié. Ce critère est inspiré de la relation (6.38) où on a posé $\mu'_k = \mu_k ||h_{k--}||$. Comme on le verra au cours des démonstrations, l'indice (k--) dans (6.42) provient de ce que $(k--)+1 \leq k-1$ et que $k-1$ est l'indice intervenant dans l'opérateur de restauration en x_k: $A(y_{k-1})^-$. A l'étape 11, l'adaptation de μ_k sert à ce que, après un nombre fini d'itérations, on ait soit $(\gamma_k, \sigma_k) > 0$, soit $\sigma_k = 0$ lorsque le critère (6.42) est vérifié. On montrera que (μ_k) est bornée inférieurement par une constante positive.

Pour l'algorithme ALG2, on désignera par K_0, l'ensemble des indices pour lesquels le critère (6.42) est vérifié:

$$K_0 := \{ k \in N : \rho_k ||r_{k+1}|| \leq \mu_k ||h_{k--}|| \, ||q_k|| \} \tag{6.43}$$

Alors $K = K_0 \wedge \{ k \in N : (\gamma_k, \sigma_k) > 0 \}$.

VI.4 - Convergence locale.

Le lemme suivant s'attache à montrer, en s'aidant du lemme VI.2.2, qu'une inégalité du type (6.12) est vérifiée pour les deux algorithmes ALG1 et ALG2 lorsque l'on pose $\delta_k := |||H_k - H^*|||$, où $|||\cdot|||$ est la norme matricielle définie en (6.6) avec $M := (G^*)^{1/2}$. Pour un indice k, on notera, en fonction de l'algorithme considéré:

$$p(k) := \begin{cases} k & \text{pour ALG1} \\ k-- & \text{pour ALG2} \end{cases}$$

On a toujours $p(k) \geq 0$.

VI.4.1 - <u>Lemme.</u>

SI . (H_0'), (H_1), (H_2') et (H_3)

. $\zeta > 0$ tel que $B(x^*,\zeta) \subset \Omega$

. k un indice

ALORS il existe des constantes positives ω_1, ω_2, ω_3, ω_4 et ε dépendant de ζ telles que si les points $(x_i : 0 \leq i \leq k+1)$ et $(y_i : 0 \leq i \leq k)$ ont été générés par l'un des algorithmes ALG1 ou ALG2 avec $y_{p(k)}$, $x_{p(k)+1}$ dans $B(x^*,\zeta)$ et y_k, x_{k+1} dans $B(x^*,\varepsilon)$, alors

$$||\gamma_k - G^* \sigma_k|| \leq \omega_1 \chi_k ||\sigma_k|| \text{ pour } k \in K_0 \tag{6.44}$$

si de plus, $y_{p(k)}$ et $x_{p(k)+1}$ sont dans $B(x^*,\varepsilon)$ alors on a

$$||\sigma_k - H^* \gamma_k|| \leq \omega_2 \chi_k ||\gamma_k|| \text{ pour } k \in K_0 \tag{6.45}$$

$$\delta_{k+1} \leq (1 - \tau_k \theta_k^2 + \omega_3 \tau_k \chi_k) \delta_k + \omega_4 \tau_k \chi_k \tag{6.46}$$

où on a noté

$$\delta_k := |||H_k - H^*||| \tag{6.47}$$

$$\tau_k := \begin{cases} 1 & \text{si } k \in K \\ 0 & \text{sinon} \end{cases} \tag{6.48}$$

$$\theta_k := \begin{cases} \dfrac{(3)^{1/2}}{4\ ||H^*||} \ \dfrac{||(H_k-H^*)\ \gamma_k||}{||H_k-H^*||\ ||\gamma_k||} & \text{si } \gamma_k \neq 0,\ H_k \neq H^* \\ 0 & \text{sinon} \end{cases} \tag{6.49}$$

$$\chi_k := \begin{cases} ||y_k-x^*||+||x_{k+1}-x^*|| & \text{pour ALG1} \\ \mu_k ||h_{k--}||+||y_k-x^*||+||x_{k+1}-x^*|| & \text{pour ALG2} \end{cases} \tag{6.50}$$

Dans (6.44) et (6.45), $K_0 = \mathbb{N}$ pour ALG1 et est donné par (6.43) pour ALG2. Dans (6.46) et (6.49), γ_k et σ_k sont donnés par (6.19)-(6.20) ou (6.26)-(6.27) selon l'algorithme considéré. La relation (6.46) découle de (6.45) et du lemme VI.2.2. La relation (6.45) se déduit de (6.44) lorsque χ_k est assez petit. Tandis que (6.44) provient des arguments utilisés au paragraphe VI.3 que l'on reprend ici mais sans supposer la convergence de (y_k) vers x^*.

Preuve.

Si $k \in K^C$, il n'y a pas de mise à jour de H_k, $H_{k+1} = H_k$, $\tau_k = 0$ et (6.46) est trivialement vérifiée. Il suffit donc de considérer les cas où $k \in K_0$. On désigne par C, C_1, C_2, ... toutes constantes dépendant de ζ et de Ω mais non des points x_i ou y_i dans Ω.

1) Comme $q_k = Z(y_k)^- Z(y_k) q_k$, on a pour l'algorithme ALG1:

$$||q_k - Z^- \sigma_k|| \le C_1 \, (||y_k - x^*|| + ||x_{k+1} - x^*||) \, ||q_k|| \tag{6.51}$$

Si on prend $\varepsilon < (2C_1)^{-1}$, on obtient:

$$||q_k|| \le ||Z^-|| \, ||\sigma_k|| \; / \; (1 - 2\varepsilon C_1) \le C(\varepsilon) \, ||\sigma_k|| \tag{6.52}$$

où $C(\varepsilon)$ décroît avec ε. D'autre part, en utilisant (2.31) sur (6.19) et en notant que g' est lipschitzienne sur Ω (hypothèse (H_0')), on obtient

$$||\tau_k - Z^{-T} L^* \, q_k|| \le C \, (||y_k - x^*|| + ||x_{k+1} - x^*||) \, ||q_k||$$

D'où l'on déduit, en tenant compte de (6.51) et (6.52):

$$||\tau_k - G^* \, \sigma_k|| \le C_2(\varepsilon) \, (||y_k - x^*|| + ||x_{k+1} - x^*||) \, ||\sigma_k|| \tag{6.53}$$

où $C_2(\varepsilon)$ dépend de ζ et décroît avec ε. Cette relation démontre (6.44) pour ALG1. En diminuant ε pour que $2\varepsilon C_2(\varepsilon)||H^*|| < 1$, on obtient:

$$||\sigma_k - H^* \, \tau_k|| \le C_3(\varepsilon) \, (||y_k - x^*|| + ||x_{k+1} - x^*||) \, ||\tau_k|| \tag{6.54}$$

où $C_3(\varepsilon) := C_2(\varepsilon)||H^*||^2 (1 - 2\varepsilon C_2(\varepsilon)||H^*||)^{-1}$. Ce qui démontre (6.45) pour ALG1.

2) Pour $k \in K_0$, le critère (6.42) de l'algorithme ALG2 est vérifié, c'est-à-dire que l'on a

$$\rho_k \, ||r_{k+1}|| \le \mu_k \, ||h_{k-}|| \, ||q_k|| \tag{6.55}$$

Comme $y_{k+1} = y_k + \rho_k \, (x_{k+1} - y_k) - \rho_k^2 \, A(y_k)^- \, c(x_{k+1})$ et $0 < \rho_k \le 1$, on peut écrire:

$$||y_{k+1} - x^*|| \le C \, (||y_k - x^*|| + ||x_{k+1} - x^*||) \tag{6.56}$$

Comme $\rho_k q_k = Z(y_k)^- Z(y_k) v_k$ et $\sigma_k = Z(y_{k+1}) v_k$, on obtient avec (6.56)

$$||\rho_k\, q_k - Z^- \sigma_k|| \le C\, (||y_k - x^*|| + ||x_{k+1} - x^*||)\, ||v_k|| \tag{6.57}$$

Comme y_{k--} et $x_{(k--)+1}$ sont dans $B(x^*, \zeta)$, on a $||h_{k--}|| \le C_4$. $v_k = \rho_k\, q_k + \rho_k^2\, r_{k+1}$ et avec (6.55), puis (6.57) et (6.56), on obtient

$$\begin{aligned} ||v_k|| &\le (1+\mu_0 C_4)\, \rho_k\, ||q_k|| \\ &\le C_5\, ||\sigma_k|| + C_6\, (||y_k - x^*|| + ||x_{k+1} - x^*||)\, ||v_k|| \end{aligned}$$

En diminuant ε de telle sorte que $\varepsilon < 1/(2C_6)$, on obtient

$$||v_k|| \le C_5\, ||\sigma_k|| \,/\, (1 - 2\varepsilon C_6) \le C(\varepsilon)\, ||\sigma_k|| \tag{6.58}$$

où, $C(\varepsilon)$ décroît avec ε. En utilisant (2.31), le caractère lipschitzien de g' sur Ω et (6.56), (6.26) donne

$$||\gamma_k - Z^{-T} L^*\, v_k|| \le C\, (||y_k - x^*|| + ||x_{k+1} - x^*||)\, ||v_k||$$

Finalement, en utilisant $v_k = \rho_k\, q_k + \rho_k^2\, r_{k+1}$, (6.57), (6.55) et (6.58), cette inégalité devient:

$$||\gamma_k - G^*\, \sigma_k|| \le C_7(\varepsilon)\, (\mu_k ||h_{k--}|| + ||y_k - x^*|| + ||x_{k+1} - x^*||)\, ||\sigma_k|| \tag{6.59}$$

où $C_7(\varepsilon)$ décroît avec ε. Cette inégalité démontre (6.44) pour ALG2. Si y_{k--} et $x_{(k--)+1}$ sont dans $B(x^*, \varepsilon)$, il existe une constante C_8 telle que $||h_{k--}|| \le \varepsilon C_8$. On obtient, comme pour passer de (6.53) à (6.54), en prenant ε assez petit tel que $(2+\mu_0 C_8)||H^*||\varepsilon C_7(\varepsilon) < 1$:

$$||\sigma_k - H^*\, \gamma_k|| \le C_9(\varepsilon)\, (\mu_k ||h_{k--}|| + ||y_k - x^*|| + ||x_{k+1} - x^*||)\, ||\gamma_k|| \tag{6.60}$$

Cette inégalité démontre (6.45) pour ALG2.

3) Pour démontrer (6.46), on utilise le lemme VI.2.2 avec $M := (G^*)^{1/2}$, la racine carrée définie positive de G^* et $A := H^*$. Si ε est assez petit pour que $\omega_2 \chi_k ||G^*|| \le 1/3$, on voit que (6.7) est vérifiée. Alors l'inégalité (6.8) conduit à (6.46) avec $\omega_3 := 4||G^*||\omega_2$. et $\omega_4 := 2\, (1+2(n-m)^{1/2})\, ||(G^*)^{1/2}||_F\, ||(G^*)^{1/2}||\, \omega_2$. ●●●

Le théorème suivant montre que les méthodes de mise à jour de H_k dans les algorithmes ALG1 et ALG2 permettent d'obtenir des suites (x_k) convergeant

Q-linéairement vers x^* si le couple initial (x_0, H_0) est choisi suffisamment proche de (x^*, H^*) (théorème local). On supposera dans ce théorème que $\rho_k = 1$ pour tout indice k. On sait en effet que lorsque le pas ρ_k diffère de 1, x_{k+1} n'est plus défini à partir du seul point x_k et cela complique singulièrement l'analyse. Toutefois, cette hypothèse sur le pas n'est pas l'hypothèse la plus contraignante. Ce qui est contraignant, c'est de demander que $||H_0 - H^*||$ et $||x_0-x^*||$ soient petits. En effet, il devrait être clair, d'après le résultat (ii) du théorème V.4.1, que le pas vaut 1 si $||x_0-x^*||$ et $||H_0-H^*||$ sont pris assez petits. Quant à la suite (y_k), on sait qu'elle ne converge en général pas Q-linéairement même si $\rho_k = 1$ et $H_k = H^*$ (c.f.r. l'exemple de Byrd (1985)).

VI.4.2 - Théorème.

SI . (H_0'), (H_1), (H_2') et (H_3)

. $0 < r < 1$, un réel

. s et t deux réels positifs

ALORS il existe des constantes positives ε et δ telles que

si $x_0 \in \Omega$ et H_0 vérifient

$$||x_0 - x^*|| < \varepsilon \tag{6.62}$$

$$||H_0 - H^*|| < \delta \tag{6.63}$$

alors, les algorithmes ALG1 ou ALG2 avec le pas unité ($\rho_k = 1$ pour tout k) génèrent des suites $(x_k) \subset \Omega$, (R_k) et (H_k) telles que

(i) (G_k) et (H_k) sont bornées

(ii) $||R_k - A^-|| < s$ pour tout indice k

(iii) $||H_k - H^*|| < t$ pour tout indice k

(iv) x_k converge Q-linéairement vers x^*:

$$||x_{k+1} - x^*|| \leq r\ ||x_k - x^*|| \text{ pour tout indice k}$$

Dans l'énoncé du théorème, $R_k = A(x_k)^-$ ou $A(y_{k-1})^-$ selon l'algorithme utilisé. L'idée de la démonstration est celle utilisée pour démontrer le théorème de détérioration bornée pour les méthodes à métrique variable en optimisation sans contrainte. On montre que l'écart $||H_k - H^*||$ ne se détériore pas trop au cours des itérations et qu'il est borné par une constante contrôlable par l'écart initial $||H_0 - H^*||$.

Preuve.
Soit $|||.|||$ la norme matricielle introduite en (6.6), avec $M := (G^*)^{1/2}$. Soit η une constante positive telle que pour tout Q dans $L(\mathbb{R}^{n-m})$, on ait:

$$\frac{1}{\eta} |||Q||| \leq ||Q|| \leq \eta \; |||Q||| \tag{6.64}$$

Soient $r \in]0,1[$, $s > 0$, $t > 0$ et $\varepsilon_0(r)$, $\beta_0(r)$, $\delta_0(r)$ donnés par le théorème IV.4.1. On fixe $\beta > 0$ tel que

$$\beta < \min (s,\beta_0)$$

Soient $\zeta > 0$ tel que

$$\begin{gathered} B(x^*,\zeta) \subset \Omega \\ ||A(z)^- - A^-|| \leq \beta \text{ pour tout } z \text{ dans } B(x^*,\zeta) \end{gathered} \tag{6.65}$$

Soient $\omega_1(\zeta)$, $\omega_2(\zeta)$, $\omega_3(\zeta)$, $\omega_4(\zeta)$ et $\varepsilon_1(\zeta)$ donnés par le lemme VI.4.2. On peut supposer $\varepsilon_1 \leq \zeta$. On fixe $\varepsilon_2 > 0$ tel que

$$x \in B(x^*,\varepsilon_2) \;\Rightarrow\; x - A_z^- c(x) \in B(x^*,\varepsilon_1) \tag{6.66}$$

quelque soit z dans $B(x^*,\zeta)$. Avec un pas unité, $(x_i : 0 \leq i \leq k+1) \subset B(x^*,\varepsilon_1)$ et $(y_i : 0 \leq i \leq k) \subset B(x^*,\varepsilon_1)$, on peut récrire l'inégalité (6.46) comme suit:

$$\begin{aligned} \delta_{k+1} \leq \delta_k + C_2 \; \tau_k \; (\omega_3 \; \delta_k + \omega_4) \\ (||x_{p(k)}-x^*||+||x_{p(k)+1}-x^*||+||x_k-x^*||+||x_{k+1}-x^*||) \end{aligned} \tag{6.67}$$

où C_2 dépend de ε_1 et δ_k, τ_k sont donnés en (6.47), (6.48). On n'y a pas repris le terme négatif à droite dans (6.46). On fixe $\delta > 0$, puis $\varepsilon > 0$ tels que

$$2 \eta^2 \; \delta < \min (t, ||H^*||^{-1}) \tag{6.68}$$

$$2 \eta^2 \; \delta \; ||G^*|| \; (||H^*||^{-1} - 2 \eta^2 \; \delta)^{-1} < \delta_0 \tag{6.69}$$

$$\varepsilon < \min (\varepsilon_0, \varepsilon_1, \varepsilon_2) \tag{6.70}$$

$$4\ \varepsilon\ C_2\ (2\ \eta\ \delta\ \omega_3 + \omega_4)\ (3 - 2\ r)\ /\ (1 - r) \le \eta\ \delta \tag{6.71}$$

Supposons que (x_0, H_0) vérifie (6.62) et (6.63) et montrons par récurrence que pour tout indice k, on a:

$$|||H_k - H^*||| \le 2\ \eta\ \delta \tag{6.72}$$

$$||R_k - A^-|| \le \beta \tag{6.73}$$

$$||x_{k+1} - x^*|| \le r\ ||x_k - x^*|| \tag{6.74}$$

$$\delta_{k+1} \le \delta_k + 4\ C_2\ \tau_k\ (2\ \eta\ \delta\ \omega_3 + \omega_4)\ ||x_{p(k)} - x^*|| \tag{6.75}$$

Remarquons que lorsque (6.72) est vérifiée, on a, d'après (6.64) et (6.68), $||H_k - H^*|| < ||H^*||^{-1}$ et donc H_k est non singulière avec $||G_k|| \le (||H^*||^{-1} - 2\eta^2\delta)^{-1}$. En utlisant cette inégalité et (6.69), on voit que (6.72) implique que

$$||G_k - G^*|| \le ||G^*||\ ||G_k||\ ||H_k - H^*|| < \delta_0 \tag{6.76}$$

Ce qui permet d'utiliser le théorème IV.4.1.

Montrons (6.72)-(6.75) pour k=0. D'après (6.63) et (6.64), $(6.72)_0$ est vérifiée. Comme $R_0 = A(x_0)^-$ et $x_0 \in B(x^*, \zeta)$, (6.65) montre que l'on a $(6.73)_0$. $(6.76)_0$, $(6.73)_0$ et $x_0 \in B(x^*, \varepsilon_0)$ permettent d'appliquer le théorème IV.4.1 qui donne $(6.74)_0$. Quant à $(6.75)_0$, il suffit d'utiliser $(6.74)_0$, $(6.72)_0$ et la conclusion (6.67) du lemme VI.4.2 puisque x_0, y_0 et x_1 sont dans $B(x^*, \varepsilon_1)$.

Supposons à présent que $(6.72)_k$, $(6.73)_k$, $(6.74)_k$, et $(6.75)_k$ soient vérifiées pour k=0,...,m-1 et montrons qu'elles le sont pour k=m. En sommant les inégalités $(6.75)_k$ pour k=0,...,m-1 et en tenant compte de (6.63), puis de (6.74) et de (6.62), et enfin de (6.71), on obtient:

$$\begin{aligned} |||H_m - H^*||| &\le |||H_0 - H^*||| + 4\ C_2\ (2\eta\delta\omega_3 + \omega_4) \sum_{k=0}^{m-1} \tau_k\ ||x_{p(k)} - x^*|| \\ &\le \eta\ \delta + 4\ C_2\ (2\eta\delta\omega_3 + \omega_4)\ \left(\sum_{k=0}^{m-1} ||x_k - x^*|| + 2\ \varepsilon\right) \\ &\le \eta\ \delta + 4\ C_2\ (2\eta\delta\omega_3 + \omega_4)\ \left(\frac{1}{1 - r} + 2\right)\ \varepsilon \\ &\le 2\ \eta\ \delta. \end{aligned}$$

On a donc $(6.72)_m$. D'après (6.74), x_{m-1} et $x_m \in B(x^*,\varepsilon)$ et d'après (6.66), y_{m-1} et $y_m \in B(x^*,\varepsilon_1)$. Alors (6.65) montre que l'on a $(6.73)_m$. On peut appliquer le théorème IV.4.1 puisque l'on a $(6.76)_m$, $(6.73)_m$ et $x_m \in B(x^*,\varepsilon_0)$. On obtient $(6.74)_m$. On est également dans les conditions requises pour appliquer le lemme VI.4.2, ce qui donne (6.67) avec $k = m$. Cette inégalité et (6.74) montrent que l'on a $(6.75)_m$.

Avec (6.72)-(6.74), on a montré (ii)-(iv). Quant au résultat (i), il s'obtient grâce à (6.72) et (6.76). ●●●

Sous les conditions du théorème précédent, on peut obtenir des suites (x_k) et (H_k) telles que (x_k) converge Q-linéairement vers x^* et (H_k) soit bornée. Alors (y_k) converge vers x^* et il existe deux constantes C_1 et C_2 indépendantes de k telles que

$$||y_{k+2} - x^*|| \le C_1 \, ||x_{k+2} - x^*||$$

$$||x_{k+2} - x^*|| \le r \, ||x_{k+1} - x^*||$$

$$||x_{k+1} - x^*|| \le C_2 \, ||y_k - x^*||$$

Dès lors,

$$||y_{k+2} - x^*|| \le C_1 \, C_2 \, r \, ||y_k - x^*||$$

Si on peut s'arranger pour que $r < (C_1C_2)^{-1}$ (C_1 et C_2 dépendent de r), la suite (y_k) convergera Q-linéairement en deux pas. C'est l'objet du corollaire suivant.

VI.4.3 - Corollaire.

SI . (H_0'), (H_1), (H_2') et (H_3)
. $0 < r < 1$, un réel
. s et t deux réels positifs

ALORS il existe des constantes positives ε et δ telles que si (x_0,H_0) vérifie (6.62) et (6.63) alors les algorithmes ALG1 ou ALG2 avec du pas unité génèrent des suites (R_k)

et (H_k) avec les conditions (i) à (iii) du théorème VI.4.2 et une suite (y_k) convergeant Q-linéairement en 2 pas vers x^*, c'est-à-dire que pour tout indice k, on a

$$||y_{k+2} - x^*|| \le r\ ||y_k - x^*||$$

Preuve.

Soient $r \in]0,1[$, $t > 0$ et $s > 0$. Soit $\zeta > 0$ tel que $B(x^*,\zeta) \subset \Omega$. Soient C_1, C_2 et C_3 des constantes telles que pour tout x dans $B(x^*,\zeta)$, on ait

$$||c'(x)|| \le C_1 \tag{6.77}$$

$$||g'(x)|| \le C_2 \tag{6.78}$$

$$||Z(x)^-|| \le C_3 \tag{6.79}$$

Soit $r_0 > 0$ tel que

$$[1 + (||A||^- + s)\ C_1]\ r_0\ [1 + (||H^*|| + t)\ C_2\ C_3] \le r \tag{6.80}$$

Soient enfin $\varepsilon_0(r_0,s,t)$ et $\delta_0(r_0,s,t)$ donnés par le théorème VI.4.2. On choisit

$$\varepsilon < \min\ (\varepsilon_0,\ \zeta,\ \zeta/[1+(||A^-||+s)C_1]) \tag{6.81}$$

$$\delta = \delta_0$$

Supposons que (x_0,H_0) vérifie les inégalités (6.62) et (6.63). Alors, d'après le théorème VI.4.2, on a pour tout indice k:

$$x_k \in B(x^*,\varepsilon) \tag{6.82}$$

$$||R_k - A^-|| < s \tag{6.83}$$

$$||H_k - H^*|| < t \tag{6.84}$$

$$||x_{k+1} - x^*|| \le r_0\ ||x_k - x^*|| \tag{6.85}$$

On a aussi $(y_k) \subset B(x^*,\zeta)$, puisque, en développant $c(x_k)$ autour de x^* et en utilisant (6.83), (6.82) et (6.77); puis (6.82) et (6.81), on trouve:

$$\begin{aligned} ||y_k - x^*|| &= ||x_k - x^* - R_k\ c(x_k)|| \\ &\le [1 + (||A^-||+s)\ C_1]\ ||x_k - x^*|| \\ &< \zeta \end{aligned}$$

On calcule, en utilisant (6.83), (6.82) et (6.77); puis (6.79), (6.76) (grâce à (6.84)), le caractère lipschitzien de g' et (6.78):

$$||r_{k+2}|| \le (||A^-|| + s)\ c_1\ ||x_{k+2} - x^*||$$

$$||y_{k+2} - x^*|| \le [1 + (||A^-|| + s)\ c_1]\ ||x_{k+2} - x^*|| \qquad (6.86)$$

$$||q_k|| \le c_3\ (||H^*|| + t)\ c_2\ ||y_k - x^*||$$

$$||x_{k+1} - x^*|| \le [1 + (||H^*|| + t)\ c_2\ c_3]\ ||y_k - x^*|| \qquad (6.87)$$

Finalement (6.86), (6.85), (6.87) et (6.80) conduisent au résultat. ●●●

VI.5 - Convergence Q-superlinéaire.

Le théorème suivant donne des conditions assurant la convergence Q-superlinéaire de (x_k) générée par ALG1.

VI.5.1 - Théorème.

SI . (H_0'), (H_1), (H_2') et (H_3)

. $(x_k) \subset \Omega$, $(y_k) \subset \Omega$ et (H_k) générées par l'algorithme ALG1

. $\sum_{k=0}^{\infty} ||x_k - x^*|| < +\infty$ (6.88)

. $\sum_{k=0}^{\infty} ||y_k - x^*|| < +\infty$ (6.89)

ALORS (i) (H_k) est bornée et uniformément définie positive

(ii) $(G_k - G^*)\ \sigma_k = o(||\sigma_k||)$ pour $k \in K$

(iii) $\rho_k = 1$ pour k assez grand

(iv) $x_k \longrightarrow x^*$ Q-superlinéairement

(v) $y_k \longrightarrow x^*$ Q-superlinéairement en 2 pas

Remarques.

Lorsque $\rho_k = 1$, les relations (6.88) et (6.89) sont équivalentes. En effet, dans ce cas $y_k := x_k - A(x_k)^- c(x_k)$ et si (x_k) converge vers x^*, on voit que $y_k - x^*$

$= o(||x_k - x^*||)$. Dès lors, (6.88) implique (6.89). Inversément, si (x_k) converge vers x^*, on a $Z(y_k)^- Z(y_k)h_k = q_k + o(||q_k||)$ et si (y_k) converge vers x^*, on en déduit que $q_k = O(||h_k||)$ et donc que $x_{k+1} - x^* = o(||y_k - x^*|| + ||h_k||)$. Comme $||h_k|| \leq ||y_{k+1} - x^*|| + ||y_k - x^*||$ $(\rho_k = 1)$, on voit que (6.89) implique (6.88). L'hypothèse (6.88) est vérifiée lorsque (x_k) converge Q-linéairement vers x^*, ce qui est assuré sous les conditions du théorème VI.4.2. L'hypothèse (6.88) est vérifiée si (y_k) converge Q-linéairement vers x^* (une situation rare), ou si la convergence est Q-linéaire en deux pas, puisqu'alors, on a

$$\sum_{k=0}^{\infty} ||y_k - x^*|| = \sum_{k=0}^{\infty} (||y_{2k} - x^*|| + ||y_{2k+1} - x^*||)$$
$$\leq \frac{1}{1 - r} (||y_0 - x^*|| + ||y_1 - x^*||)$$

où, $0 < r < 1$. La convergence Q-linéaire en 2 pas de (y_k) est assurée sous les conditions du corollaire VI.4.3.

La preuve du théorème se déroule comme suit. La clause (i) provient du théorème de Powell (VI.2.3). Pour (ii), on utilise la technique de Broyden, Dennis, Moré (1973). Le théorème V.4.1 et (ii) permettent de montrer que le pas vaut 1 après un nombre fini d'itérations. Alors, asymptotiquement, on se retrouve dans le cadre de l'algorithme modèle local de la section IV. Puisque (x_k) converge vers x^*, $R_k \longrightarrow A^-$ et (H_5) est vérifiée. (H_6) est également vérifiée grâce à (i). On peut alors utiliser la caractérisation de la convergence Q-superlinéaire donnée au théorème IV.5.1 qui dit que (x_k) convergera Q-superlinéairement vers x^* SSI (c.f.r. (4.67))

$$(G_k - G^*) \frac{Z\, d_k}{||x_k - x^*||} \longrightarrow \infty \quad \text{quand} \quad k \longrightarrow \infty \tag{6.90}$$

Pour $k \in K$, cette condition est plus faible que ce que la mise à jour de H_k donnera: (ii). Mais le résultat (ii) aura aussi donné plus que la convergence Q-superlinéaire puisque c'est grâce à lui que le pas ρ_k vaudra 1 après un nombre fini d'itérations (pour $k \in K$). Enfin, (v) est une conséquence immédiate de (iv).

<u>Preuve</u>.

1) <u>On prouve (i) et (ii)</u>. Soit $\zeta > 0$ tel que $B(x^*, \zeta)$ soit contenue dans Ω. Soient $\omega_1(\zeta)$, $\omega_2(\zeta)$, $\omega_3(\zeta)$, $\omega_4(\zeta)$ et $\varepsilon(\zeta)$ donnés par le lemme VI.4.1. On peut supposer $\varepsilon \leq \zeta$. Comme (x_k) et (y_k) converge vers x^*, il existe un indice K_1 tel que pour k supérieur à K_1 on ait

$$||x_k - x^*|| < \varepsilon$$

$$||y_k - x^*|| < \varepsilon$$

Ce qui permet d'appliquer le lemme VI.4.1 dès que $k \geq K_1$. Des inégalités (6.45) et (6.46), on obtient:

$$||\sigma_k - H^* \gamma_k|| \leq \omega_2\, \chi_k\, ||\gamma_k|| \tag{6.91}$$

$$\delta_{k+1} \leq (1 - \tau_k\, \theta_k^2 + \omega_3\, \tau_k\, \chi_k)\, \delta_k + \omega_4\, \tau_k\, \chi_k \tag{6.92}$$

$$\chi_k := ||y_k - x^*|| + ||x_{k+1} - x^*||$$

où δ_k, τ_k, et θ_k sont donnés en (6.47)-(6.49). D'après (6.88) et (6.89), (χ_k) est sommable. Comme $H_{k+1} = H_k$ si $k \in K^c$, l'inégalité (6.91) et le fait que $(\gamma_k, \sigma_k) > 0$ pour $k \in K$ permettent d'appliquer le lemme VI.2.3, ce qui conduit à (i). La suite $(\tau_k \chi_k)$ est également sommable. Alors, l'inégalité (6.92) et le lemme VI.2.4 montre que (δ_k) converge, disons vers δ, et que soit $\delta = 0$, soit $\tau_k \theta_k$ converge vers 0. Dans chaque cas, on a (ii).

2) <u>On prouve (iii)</u>. D'après (6.91), $\sigma_k = H^* \gamma_k + o(||\gamma_k||)$ et donc $\gamma_k = G^* \sigma_k + o(||\sigma_k||)$. Comme G^* est définie positive (hypothèse (H_2')), il existe un indice $K_2 \geq K_1$ et une constante C_1 tels que pour tout $k \geq K_2$, on ait

$$(\gamma_k, \sigma_k) \geq C_1\, ||\sigma_k||^2$$

Dès lors, pour $k \geq K_2$, soit $\sigma_k \neq 0$, soit $k \in K$. Pour ALG1, on a $\sigma_k = Z(x_{k+1}) q_k = Z q_k + o(||q_k||)$. (G_k) étant bornée, la propriété (P_2) du théorème V.4.1 est vérifiée pour les indices k tels que $\sigma_k = 0$. Dès lors, ce théorème montre qu'il existe un indice $K_3 \geq K_2$ tel que $\rho_k = 1$ si $k \geq K_3$ et $\sigma_k = 0$. D'autre part, si on injecte la relation $\sigma_k = Z q_k + o(||q_k||)$ dans (i), on voit que la propriété (P_2) du théorème V.4.1 est à nouveau vérifiée, ce qui conduit encore à l'existence d'un indice $K_4 \geq K_2$ au delà duquel $\rho_k = 1$ si $k \in K$. Avec $K_5 := \max(K_3, K_4)$, on a $\rho_k = 1$ pour tout indice $k \geq K_5$.

3) <u>On prouve (iv)</u>. On suppose $k \geq K_5$. Comme alors $\rho_{k-1} = 1$, on peut utiliser les résultats de la section IV, notamment $q_k = O(||x_k-x^*||)$ et $d_k = O(||x_k-x^*||)$. Lorsque $\rho_{k-1} = 1$, on a aussi $\sigma_k = Zd_k + o(||d_k||)$. De cette relation, de $\sigma_k = Zq_k + o(||q_k||)$ et de (ii), on déduit:

$$Z\, d_k = o(||x_k-x^*||) \text{ pour } \sigma_k = 0 \tag{6.93}$$

$$(G_k - G^*)\, Z\, d_k = o(||x_k-x^*||) \text{ pour } k \in K \tag{6.94}$$

Comme pour $k \geq K_5$, soit $\sigma_k = 0$, soit $k \in K$, (6.93) et (6.94) montrent que (6.90) est vérifiée. Par conséquent, la suite (x_k) converge Q-superlinéairement vers x^*.

4) <u>On prouve (v)</u>. Le pas vaut 1 pour $k \geq K_5$ et (H_k) est bornée. Alors il existe deux constantes positives C_2 et C_3 telles que pour tout $k \geq K_5$ on ait

$$||y_{k+2} - x^*|| \leq C_2\, ||x_{k+2} - x^*||$$

$$||x_{k+1} - x^*|| \leq C_3\, ||y_k - x^*||$$

Comme $||x_{k+2}-x^*||/||x_{k+1}-x^*||$ converge vers 0, on voit que

$$\frac{||y_{k+2} - x^*||}{||y_k - x^*||} \longrightarrow 0$$

Ce qui montre la convergence Q-superlinéaire en 2 pas de la suite (y_k). ●●●

Le théorème suivant donne des conditions pour que la suite (x_k) converge Q-superlinéairent lorsqu'elle est générée par ALG2.

VI.5.2 - <u>Théorème.</u>

SI . (H_0'), (H_1), (H_2') et (H_3)
 . $(x_k) \subset \Omega$, $(y_k) \subset \Omega$ et (H_k) générées par l'algorithme ALG2
 . (6.88) et (6.89)
 . $$C_1 := \sup_{k\geq 0} \sup_{l\geq 0} \frac{||y_{k+1} - x^*||}{||y_k - x^*||} < +\infty \tag{6.95}$$

ALORS	(i)	(H_k) bornée et uniformément définie positive
	(ii)	$(G_k - G^*)\ \sigma_k = o(\|\|\sigma_k\|\|)$ pour $k \in K$
	(iii)	$y_{k+1} - x^* = o(\|\|y_k - x^*\|\|)$ pour $k \in K$
	(iv)	$\rho_k = 1$ pour k assez grand
	(v)	$x_k \longrightarrow x^*$ Q-superlinéairement
	(vi)	$y_k \longrightarrow x^*$ Q-superlinéairement en 2 pas

Remarques.

Par rapport aux hypothèses du théorème VI.5.1, on a rajouté (6.95). Cette hypothèse vient de ce qu'on a besoin de comparer $||y_k - x^*||$ à $||y_{k--} - x^*||$ et que $(k-(k--))$ n'est pas nécessairement bornée. (6.95) est vérifiée par exemple sous les conditions du corollaire VI.4.3. En effet, dans ce cas, (H_k) étant bornée, il existe une constante C, indépendante de k, telle que pour tout indice k on ait

$$||y_{k+1} - x^*|| \leq C\ ||y_k - x^*||$$

Alors quelque soit k et l, la convergence Q-linéaire en 2 pas de (y_k) montre que l'on a

$$||y_{k+1} - x^*|| \leq C\ ||y_k - x^*||$$

Et par conséquent (6.95) est vérifiée. Dans (ii), σ_k est donné cette fois par (6.27).

Preuve.

1) On prouve (i) et (ii). Soit $\zeta > 0$ tel que $B(x^*, \zeta)$ soit contenue dans Ω. Soient $\omega_1(\zeta)$, $\omega_2(\zeta)$, $\omega_3(\zeta)$, $\omega_4(\zeta)$ et $\varepsilon(\zeta)$ donnés par le lemme VI.4.1. On peut supposer $\varepsilon \leq \zeta$. Comme (x_k) et (y_k) convergent vers x^*, il existe un indice K_1 tel que x_k et $y_k \in B(x^*, \varepsilon)$ pour $k \geq K_1$. L'application du lemme VI.4.1 donne:

$$||\gamma_k - G^*\ \sigma_k|| \leq \omega_1\ \chi_k\ ||\sigma_k|| \quad \text{pour } k \in K_0,\ k \geq K_1 \tag{6.96}$$

$$\chi_k := \mu_k\ ||h_{k--}|| + ||y_k - x^*|| + ||x_{k+1} - x^*|| \tag{6.97}$$

Si K est borné, (i) est clairement vérifiée et pour (ii) il n'y a rien à démontrer. Si K n'est pas borné, on peut trouver un indice $K_2 \geq K_1$ tel que $(k--) \geq K_1$ si $k \geq K_2$. L'application du lemme VI.4.1 pour $k \geq K_2$ donne:

$$||\sigma_k - H^* \tau_k|| \leq \omega_2 \chi_k ||\tau_k|| \quad \text{pour } k \in K_0 \text{ (K non borné)} \tag{6.98}$$

$$\delta_{k+1} \leq (1 - \tau_k \theta_k^2 + \omega_3 \tau_k \chi_k) \delta_k + \omega_4 \tau_k \chi_k \text{ (K non borné)} \tag{6.99}$$

où δ_k, τ_k, et θ_k sont donnés en (6.47)-(6.49). Comme (R_k) est bornée, il existe une constante C_2 telle que $||r_{k+1}|| \leq C_2||x_{k+1}-x^*||$. On en déduit que pour tout indice k, on a:

$$||h_k|| \leq (1 + C_2) ||x_{k+1} - x^*|| + ||y_k - x^*|| \tag{6.100}$$

Alors, on peut majorer χ_k par

$$\chi_k \leq (1+C_2)\mu_k||x_{(k--)+1}-x^*|| + \mu_k||y_{k--}-x^*|| + ||y_k-x^*|| + ||x_{k+1}-x^*||$$

Grâce à cette inégalité, (6.88) et (6.89), on voit que $(\tau_k\chi_k)$ est sommable. En effet, (μ_k) est bornée par μ_0 et, par exemple, $(\tau_k||y_{k--}-x^*||)$ est sommable grâce à (6.89) et au fait que parmi les éléments non nuls de cette suite, chaque écart $||y_i-x^*||$, $i > 1$ apparaît au plus une fois. Pour la sous-suite d'indices K, c'est-à-dire chaque fois que l'on modifie H_k, (6.98) montre que l'on a (6.11). On a également (6.10) pour $k \in K$. Dès lors, d'après le lemme VI.2.3, (i) est vérifiée. Grâce à l'inégalité (6.99), on peut appliquer le lemme VI.2.4 qui montre que (δ_k) converge, disons vers δ, et que soit $\delta = 0$, soit $\tau_k\theta_k$ converge vers 0. Dans chaque cas, on a (ii).

2) <u>Deux inégalités lorsque $\rho_{k-1} = 1$</u>. On désignera par C une constante positive "générique" indépendante de k et par (η_k) une suite "générique" de réels positifs convergeant vers 0 lorsque k tend vers l'infini. Soit k un indice tel que $\rho_{k-1} = 1$. Dans ce cas on peut utiliser certaines inégalités de la section IV. Comme on a (6.1), de (4.62) (équivalent à (4.64)) et de (4.76), on obtient:

$$||x_{k+1} - x^*|| \leq ||Z^- Z (x_k - x^*) + q_k|| + \eta_k ||x_k - x^*|| \tag{6.101}$$

Comme (H_k) est bornée, les relations (4.33) et (4.31) donnent:

$$q_k = - Z^- H_k G^* Z (x_k - x^*) + o(||x_k - x^*||) \tag{6.102}$$

Grâce à l'uniforme non singularité de (H_k), on en déduit

$$||Z (x_k - x^*)|| \le C ||q_k|| + \eta_k ||x_k - x^*|| \tag{6.103}$$

On a également en utilisant (6.102):

$$Z^- Z (x_k - x^*) + q_k = Z^- H_k (G_k - G^*) Z (x_k - x^*) + o(||x_k - x^*||) \tag{6.104}$$

Alors, la relation (6.101) devient, en tenant compte de (6.104), (H_k) bornée et (6.103):

$$||x_{k+1} - x^*|| \le C_3 ||G_k - G^*|| \, ||q_k|| + \eta_k ||x_k - x^*|| \tag{6.105}$$

D'autre part, d'après (4.36) et (4.31), on a

$$r_{k+1} = A^- A (A(y_{k-1})^- - A^-) A (x_k - x^*) + O(||x_k - x^*||^2)$$

Comme $x_k - x^* = O(||y_{k-1} - x^*||)$ $(\rho_{k-1} = 1)$, on obtient:

$$||r_{k+1}|| \le C_4 ||y_{k-1} - x^*|| \, ||x_k - x^*|| \tag{6.106}$$

3) Pas unité pour k assez grand dans K. De la définition (6.27) de σ_k ou de (6.57), on obtient

$$\sigma_k = \rho_k Z q_k + \rho_k o(||q_k||) \tag{6.107}$$
$$\sigma_k = \rho_k O(||q_k||)$$

Ces estimations, (G_k) bornée et (ii) donnent:

$$(G_k - G^*) Z q_k = o(||q_k||) \text{ pour } k \in K \tag{6.108}$$

C'est la propriété (P_2) du théorème V.4.1 qui montre que $\rho_k = 1$ après un nombre fini d'itérations, disons K_3.

4) On prouve (iii). Si K est borné, il n'y a rien à démontrer. On suppose donc que K n'est pas borné. Pour $k \in K$ et $k \ge K_3$, le critère (6.42) est satisfait. Comme $\rho_k = 1$, on trouve:

$$r_{k+1} = o(||q_k||) \quad \text{d'où} \quad ||q_k|| \sim ||h_k|| \tag{6.109}$$

Comme (H_k) est bornée, $x_{k+1}-x^* = O(||y_k-x^*||)$. Alors, (4.30) en k+1, (6.1), (6.109), l'équivalence (4.41) et l'injectivité de A^- donnent:

$$A\ (x_{k+1} - x^*) = o(||h_k||) \tag{6.110}$$

D'autre part, comme on a obtenu (4.77), on peut obtenir

$$\begin{aligned} -\ (G_k - G^*)\ Z\ q_k &= g(y_k) + G^*\ Z\ q_k + o(||q_k||) \\ &= g(x_{k+1}) + o(||q_k||) \end{aligned} \tag{6.111}$$

En développant $g(x_{k+1})$ autour de x^* en tenant compte de (2.31), $x_{k+1}-x^* = O(||y_k-x^*||)$ et (6.110), on a

$$g(x_{k+1}) = G^*\ Z\ (x_{k+1} - x^*) + o(||h_k||) \tag{6.112}$$

Avec (6.108), (6.111) et (6.112), on voit que $Z\ (x_{k+1}-x^*) = o(||h_k||)$. Ce qui, joint à (6.110), (4.40) et le fait que $\rho_k = 1$ montre que $d_{k+1} = o(||h_k||)$. Avec (6.109), on a $q_{k+1} = o(||h_k||)$. Avec (6.1) et $\rho_k = 1$, on voit que $r_{k+2} = o(||d_k||)$. Finalement, on trouve:

$$h_{k+1} = o(||h_k||) \tag{6.113}$$

C'est à dire (iii).

5) <u>On montre qu'il existe un indice $K_4 \geq K_3$ tel que si $k \in K_0$ et $k \geq K_4$, alors soit $(\gamma_k,\sigma_k) > 0$, soit $\sigma_k = 0$. Donc $\lim_k \mu_k =: \mu_\infty > 0$</u>. Supposons $k \in K_0$. Lorsque K est non borné, $\chi_k \to 0$ et d'après (6.96), $\gamma_k = \bar{G}^*\sigma_k + o(||\sigma_k||)$. Dès lors, il existe une constante positive C_5 indépendante de k et un indice K_5 tels que

$$(\gamma_k,\sigma_k) \geq C_5\ ||\sigma_k||^2 \quad \text{pour } k \geq K_5 \text{ et } k \in K_0 \quad (K \text{ non borné}) \tag{6.114}$$

Cette inégalité prouve l'affirmation lorsque K est borné. Supposons à présent que K soit borné. On note:

$$D := \max\ (\ \max_k\ (||x_k - x^*||)\ ,\ \max_k\ (||y_k - x^*||)\ )$$

$$\bar{\mu} := (\ (2 + C_2)\ D\ \omega_1\ ||H^*||\ )^{-1}$$

D'après (6.100), on a $||h_k|| \leq (2+C_2)D$. Comme la plus petite valeur propre de G^* est $||H^*||^{-1}$, on obtient en utilisant (6.96), puis (6.97):

$$\begin{aligned}(\gamma_k,\sigma_k) &= (G^*\sigma_k,\sigma_k) + (\gamma_k - G^*\sigma_k,\sigma_k)\\ &\geq (||H^*||^{-1} - \omega_1\, \chi_k)\ ||\sigma_k||^2\\ &\geq (||H^*||^{-1} - \omega_1\, \mu_k\ (2 + C_2)\ D - \eta_k')\ ||\sigma_k||^2\\ &= (\bar{\mu}\ ||H^*||)^{-1}\ (\bar{\mu} - \mu_k - \eta_k)\ ||\sigma_k||^2 \end{aligned} \tag{6.115}$$

où (η_k) et (η_k') sont des suites de réels positifs convergeant vers 0. Si l'affirmation était fausse, il existerait une infinité d'indices $k \in K_0$, $k \geq K_1$, pour lesquels on aurait $(\gamma_k,\sigma_k) \leq 0$ et $\sigma_k \neq 0$. De plus, par construction de μ_k, on aurait $\mu_k \longrightarrow 0$. En prenant k assez grand dans (6.115), on obtiendrait $\sigma_k = 0$. Cette contradiction ($0 = \sigma_k \neq 0$) démontre l'affirmation.

6) <u>On prouve (iv)</u>. On a déjà montré que si $k \in K$ et $k \geq K_3$, alors $\rho_k = 1$. Considérons les indices $k \in K_0 \setminus K$, $k \geq K_4$. On a montré à l'étape 5 que pour ces indices k, $\sigma_k = 0$. D'après (6.107) et l'équivalence $||q_k|| \sim ||Zq_k||$, on a alors $q_k = o(||h_k||)$. C'est-à-dire que la propriété (P_1) ou (P_2) du théorème V.4.1 est vérifiée. Cela montre que $\rho_k = 1$ pour k assez grand dans $K_0 \setminus K$, disons supérieur à $K_6 \geq K_4$. Il reste à examiner les indices $k \in K_0^c$. Si K est borné, k-- est un indice constant dès que k est assez grand. Le critère (6.42) n'étant pas vérifié et $\mu_k \geq \mu_\infty > 0$ (étape 5), on a alors $q_k = O(||r_{k+1}||)$. Ceci montre que la sous-suite d'indices considérée vérifie la propriété (P_1) du théorème V.4.1 et par conséquent il existe un indice $K_7 \geq K_6$ tel que $\rho_k = 1$ si $k \geq K_6$, $k \in K_0^c$. Supposons à présent que K ne soit pas borné et que l'on n'ait pas $\rho_k = 1$ après un nombre fini d'itérations pour les indices $k \in K_0^c$. Comme K_0 n'est pas borné (car K ne l'est pas) et que $\rho_k = 1$ pour $k \in K_0$, $k \geq K_6$, on peut trouver une sous-suite d'indices (k_i) telle que:

$$\rho_{k_i} \neq 1,\ \rho_{k_i-1} = 1 \text{ et } k_i \in K_0^c$$

On montre que cela conduit à une contradiction. Pour ces indices k_i, le critère (6.42) n'étant pas vérifié, on a successivement, en utilisant (6.106) et l'équivalence (4.41) (d'où provient la constante C_6), puis (6.95) (car $(k_i--)+1 \leq k_i-1$):

$$||q_{k_i}|| \le \frac{1}{\mu} \; \frac{||r_{k_i+1}||}{||h_{k_i--}||} \tag{6.116}$$

$$\le \frac{C_4 C_6}{\mu} \; \frac{||y_{k_i-1} - x^*||}{||y_{k_i--} - x^*||} \; ||x_{k_i} - x^*||$$

$$\le \frac{C_1 C_4 C_6}{\mu} \; \frac{||y_{(k_i--)+1} - x^*||}{||y_{k_i--} - x^*||} \; ||x_{k_i} - x^*||$$

Comme $(k_i--) \in K$ qui est non borné, on a alors grâce à (iii) et à l'équivalence (4.40):

$$q_{k_i} = o(||d_{k_i}||) \tag{6.117}$$

La relation (6.117) et les hypothèses sur la sous-suite (k_i) montrent que celle-ci vérifie la propriété (P_4) du théorème V.4.1. Dès lors,

$$r_{k_i+1} = o(||h_{k_i}|| \; ||d_{k_i}||) \tag{6.118}$$

Le pas d'indice k_i-1 valant 1, on a grâce à (6.95):

$$||d_{k_i}|| \le C_7 \; ||y_{k_i-1} - x^*|| \le C_1 \; C_7 \; ||y_{k_i--} - x^*||$$

L'utilisation de cette inégalité et de (6.118) à la place de (6.106) en (6.116) conduit à:

$$q_{k_i} = o(||h_{k_i}||)$$

C'est-à-dire la propriété (P_1) du théorème V.4.1. Il y a alors contradiction entre la conclusion de ce théorème et le fait que le pas de la sous-suite considérée diffère de 1. En conclusion, on a montré l'existence d'un indice K_8 tel que

$$\rho_k = 1 \quad \text{pour tout } k \ge K_8$$

7) <u>On prouve (v)</u>. On supposera $k > K_8$ de sorte que $\rho_{k-1} = \rho_k = 1$. Considérons en premier lieu les indices $k \in K$ (supposé non borné). $r_{k+1} = o(||d_k||)$ et d'après (6.113), $q_{k+1} = - r_{k+2} + o(||h_k||) = o(||d_k||)$. Dès lors,

$$d_{k+1} = o(||d_k||) \text{ pour } k \in K \tag{6.119}$$

Considérons à présent les indices $k \in K_0 \setminus K$. D'après l'étape 5, on a $\sigma_k = 0$. Comme $\rho_{k-1} = 1$, on a $\sigma_k = Zq_k + o(||d_k||)$. Avec $\sigma_k = 0$ et l'équivalence $||q_k|| \sim ||Zq_k||$, cela donne $q_k = o(||d_k||)$ et grâce à (6.105), on obtient:

$$x_{k+1} - x^* = o(||x_k - x^*||) \quad \text{pour } k \in K_0 \setminus K \tag{6.120}$$

Considérons enfin les indices $k \in K_0^c$. Lorsque K est borné, (6.42) (non vérifiée) montre que l'on a $q_k = O(||r_{k+1}||)$ ($\mu_k \geq \mu_\infty > 0$) et les inégalités (6.106) et (6.105) conduisent à

$$x_{k+1} - x^* = o(||x_k - x^*||) \quad \text{pour } k \in K_0^c \quad (K \text{ borné}) \tag{6.121}$$

Lorsque K n'est pas borné, on a vu en (6.117) que $q_k = o(||d_k||)$. Ce qui, avec (6.105), donne

$$x_{k+1} - x^* = o(||x_k - x^*||) \quad \text{pour } k \in K_0^c \quad (K \text{ non borné}) \tag{6.122}$$

(6.119), (6.120), (6.121) et (6.122) forment le résultat (v).

8) <u>On prouve (vi)</u>. Voir l'étape 4 de la preuve du théorème VI.5.2. ●●●

CONCLUSION.

Nous avons étudié en détails (étude locale, étude globale et mise à jour de la matrice réduite dans le cadre de la méthode globale) deux algorithmes à métrique variable réduite pour le problème (1.1): ALG1 et ALG2. Ces algorithmes génèrent une suite (x_k) convergeant Q-superlinéairement vers x^* satisfaisant les conditions suffisantes d'optimalité du problème (1.1) et une suite (H_k) de matrices d'ordre (n-m) symétriques, définies positives. Celles-ci sont générées par la formule de mise à jour BFGS sur base d'une équation de quasi-Newton. Les deux algorithmes se distinguent par la technique de mise à jour de H_k. Avec ALG1, il est nécessaire de linéariser deux fois les contraintes par itération tandis qu'une seule linéarisation est nécessaire avec ALG2. Avec ALG1, la matrice réduite est mise à jour à chaque itération tandis qu'avec ALG2, la mise à jour de H_k est conditionnée à la réalisation d'un critère formulé en terme de l'information disponible à l'itération courante.

REFERENCES.

L. Armijo (1966). Minimization of functions having lipschitz continuous first partial derivatives, Pacific J. Maths 16/1, 1-3.

J. Blum, J.Ch. Gilbert, B. Thooris (1985). Parametric identification of the plasma current density from the magnetic measurements and the pressure profile, code IDENTC. Centre d'études nucléaires (DRFC), B.P. 6, 92265 Fontenay-aux-Roses (France).

P.T. Boggs, J.W. Tolle, P. Wang (1982). On the local convergence of quasi-Newton methods for constrained optimization, SIAM J. Control Optim. 20/2, 161-171.

C.G. Broyden (1969). A new double-rank minimization algorithm, Notices of the American Mathematical Society 16, 670.

C.G. Broyden, J.E. Dennis, J.J. Moré (1973). On the local and superlinear convergence of quasi-Newton methods, J. Inst. Maths Applics 12, 223-245.

R.H. Byrd (1985). An example of irregular convergence in some constrained optimization methods that use the projected hessian, Math. Progr. 32, 232-237.

R.M. Chamberlain, C. Lemarechal, H.C. Pedersen, M.J.D. Powell (1982). The watchdog technique for forcing convergence in algorithms for constrained optimization, Math. Progr. Study 16, 1-17.

P.G. Ciarlet (1982). Introduction à l'analyse numérique matricielle et à l'optimisation, Masson, Paris.

T.F. Coleman, A.R. Conn (1982,a). Nonlinear programming via an exact penalty function: asymptotic analysis, Math. Progr. 24, 123-136.

T.F. Coleman, A.R. Conn (1982,b). Nonlinear programming via an exact penalty function: global analysis, Math. Progr. 24, 137-161.

T.F. Coleman, A.R. Conn (1984). On the local convergence of a quasi-Newton method for the nonlinear programming problem, SIAM J. Num. Anal. 21/4, 755-769.

J.E. Dennis, J.J. Moré (1974). A characterization of superlinear convergence and its applications to quasi-Newton methods, Math. Comp. 28/126, 549-560.

J.E. Dennis, J.J. Moré (1977). Quasi-Newton methods, motivation and theory, SIAM Review 19, 46-89.

R. Fletcher (1970). A new approach to variable metric algorithms, Computer J. 13, 317-322.

R. Fletcher (1981). Practical methods of optimization, vol 2: constrained optimization, John Wiley & Sons (Chichester).

D. Gabay (1982,a). Minimizing a differentiable function over a differentiable manifold, J. Opt. Theory Appl. 37/2, 171-219.

D. Gabay (1982,b). Reduced quasi-Newton methods with feasibility improvement for nonlinearly constrained optimization, Math. Progr. Study 16, 18-44.

U.M. Garcia Palomares, O.L. Mangasarian (1976). Superlinearly convergent quasi-Newton algorithms for nonlinearly constrained optimization problems, Math. Prog. 11, 1-13.

J.Ch. Gilbert. Thèse, Université de Paris VI (à paraître).

D. Goldfarb (1970). A family of variable metric methods derived by variational means, Math. of Comp. 24, 23-26.

S.P. Han (1976). Superlinearly convergent variable metric algorithms for general nonlinear programming problems, Math. Progr. 11, 263-282.

S.P. Han (1977). A globally convergent method for nonlinear programming, J. Opt. Theory Appl. 22/3, 297-309.

D.Q. Mayne, E. Polak (1982). A superlinearly convergent algorithm for constrained optimization problems, Math. Progr. Study 16, 45-61.

N. Maratos (1978). Exact penalty function algorithms for finite dimensional and control optimization problems, Ph.D. Thesis, University of London.

J.W. Milnor (1965). Topology from the differentiable viewpoint, University of Virgina Press, Charlottesville, V.A.

J. Nocedal, M.L. Overton (1985). Projected hessian updating algorithms for nonlinearly constrained optimization, SIAM J. Numer. Anal. 22/5, 821-850.

J.M. Ortega, W.C. Rheinboldt (1970). Iterative solution of nonlinear equations of several variables, Academic Press, New York.

M.J.D. Powell (1971). On the convergence of the variable metric algorithm, J. Inst. Maths Apllics 7, 21-36.

M.J.D. Powell (1978). The convergence of the variable metric methods for nonlinearly constrained optimization calculations, Nonlinear Programming 3, eds: O.L. Mangasarian, R.R Meyer, S.M. Robinson, Academic Press, 27-63.

D.F. Shanno (1970). Conditioning of quasi-Newton methods for function minimization, Math. of Comp. 24, 647-656.

Y. Yuan (1985). An only 2-step Q-superlinear convergence example for some algorithms that use reduced hessian approximations, Math. Progr. 32, 224-231.

Imprimé en France

par

l'Institut National de Recherche en Informatique et en Automatique